tanulmányok 121/1981

MTA Számitástechnikai és Automatizálási Kutató Intézet Budapest

MAGYAR TUDOMÁNYOS AKADÉMIA
SZÁMITÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZETE

VIZSGÁLATOK A KEZDETI ÉRTÉK PROBLÉMÁK NUMERIKUS MEGOLDÁSÁVAL KAPCSOLATBAN

Irta:
Vicsek Tamásné (Strehó Mária)

Tanulmányok 121/1981.

A kiadásért felelős:
DR VÁMOS TIBOR

ISBN 963 311 118 8
ISSN 0324-2951

Készült a
KSH Nemzetközi Számitástechnikai Oktató és Tájékoztató
Központ Reprográfiai Uzemében
1981/117.

TARTALOMJEGYZÉK

BEVEZETÉS

A környező világ számos jelenségét modellezhetjük differenciálegyenletek (röviden DE-k) segitségével. Egy adott modell vizsgálatához szükségünk van általában a DE megoldására. Ismeretes azonban, hogy az egyenletek nagy része nem oldható meg explicit alakban azaz nem irható fel véges számu elemi müvelet és elemi (esetleg speciális) függvény segitségével. A reálisan megoldható feladatok osztálya a közelitő módszerek fejlődésével és az elektronikus számitógépek alkalmazásával bővült ki jelentősen.

A dolgozatban a közönséges DE-kre vonatkozó Cauchy-féle kezdeti értékprobléma közelitő módszereivel kapcsolatban végeztünk vizsgálatokat.

A dolgozat első fejezetében összefoglaljuk a DE közelitő megoldására vonatkozó klasszikus eredményeket (Dahlquist (1956), Henrici (1962)), kibővitve azokat az irodalomban megjelent legfrissebb anyaggal, amely a változó lépéshosszu módszereket vizsgálja.

A II. fejezetben az egyes közelitő módszerek gyakorlati alkalmazásának kritikai leirását adjuk. Az irodalomban ismertetett eredmények alapján összehasonlitjuk őket különböző optimalitási szempontokból vizsgálva hatékonyságukat. A vizsgált módszerek közül számosnak adaptáltuk a programját az Akadémia CDC 3300-as gépére. A programokat terjedelmük miatt nem mellékeljük, megtalálhatók a CDC 3300-as gép NUMCOSY programkönyvtárában [28].

Számitógépes tapasztalataink alapján a gyakorlatban a leghatékonyabbnak és legáltalánosabban használhatónak a lineáris többlépéses módszerek közül a változó rendü és változó lépéshosszu formulákkal dolgozó Gear (1971) módszer bizonyult. Ez motiválta, hogy vizsgáljuk egylépéses módszerek esetén a változó formulával a változó lépéshosszal kapcsolatos elméleti problémákat. A gyakorlatban sokszor előfordulnak (pl. biokémiában, számelméletben, epidemics stb.) olyan DE-k amelyekben a deriváltak a megoldás korábbi értékeitől is függenek. Ilyen általános tipusu un. funkcionál DE-kel fogunk foglalkozni, amelyek közé tartoznak a közönséges DE-k is. A III. fejezetben ezek közelitő megoldására definiált általánositott egylépéses módszercsalád konvergenciáját vizsgáljuk. Ez Tavernini (1972) ekvidisztans felosztásra vonatkozó eredményének általánositása. Definiáljuk és vizsgáljuk továbbá a módszercsalád stabilitását.

Végül ezen a helyen is szeretnék köszönetet mondani Dr Móricz Ferencnek a kézirat átolvasásáért, valamint értékes megjegyzéseiért.

I. A CAUCHY - FELADAT NUMERIKUS MEGOLDÁSA

1. A FELADAT KITÜZÉSE, ALAPFOGALMAK

Tekintsük a közönséges differenciálegyenletekre (DE-kre) vonatkozó

(1.1) $y' = f(x,y)$.

(1.2) $y(x_o) = y_o$ $\qquad$ (x_o, y_o adott valós konstans)

kezdeti érték problémát (Cauchy-feladatot), ahol az

$$f: I \times R^m \to R^m \quad (I = [x_o, b] \subset R^1)$$

leképezés folytonos és kielégiti az

$$\| f(x,y) - f(x,y^*) \| \leq L \| y - y^* \|$$

egyenletes Lipschitz feltételt minden $x \in I, y, y^* \in R^m$ esetén (ahol R^m az m dimenziós euklideszi tér, b és L valós konstansok és $\| \cdot \|$ pl. a maximum norma).

A továbbiakban mindig feltételezzük, hogy a fenti feltételek teljesülnek, ekkor a Picard-Lindelöf tétel szerint az (1.1), (1.2) DE-nek létezik egyértelmü megoldása (Henrici (1962)).

Keressük az (1.1) egyenletrendszernek az (1.2) kezdeti feltétel által meghatározott $y(x)$ partikuláris megoldását az $x_o \leq x \leq b$ intervallumon. A feladat közelitő megoldásakor az a célunk, hogy az $\Delta_N = \{x\ , x\ , \ldots, x_N\} \in I$ diszkrét ponthalmazon meghatározzuk az $y(x)$ partikuláris megoldás számértékeit.

Legyen π az $[x_o, b]$ intervallum összes felosztásainak halmaza. Jelöljön $\Delta_N \in \pi$ olyan felosztássorozatot, amelyre $x_N = x(N=0,1,\ldots)$. Legyen $\| \Delta_N \| = \max\limits_{0 \leq i \leq N-1} h_i$

ahol $h_i = x_{i+1} - x_i$ az i-edik lépéshossz és $\{\Delta_N\}_{N=1}^{\infty} \subset \pi$ mindig olyan, hogy teljesüljön a $\| \Delta_N \| \to 0$ $(N \to +\infty)$ feltétel.

Az x_n pontok számos esetben ekvidisztáns elhelyezésüek, azaz

$x_n = x_o + nh,\ n = 1,2,\dots,N.$

Jelölje y_n az (1.1),(1.2) feladat x_n pontbeli közeliő megoldását és $y(x_n)$ a pontos, elméleti megoldást.

Tekintsük a k-lépéses közelitő módszerek alábbi általános módszerosztályát:

$$\begin{aligned} & y_r = s_r(f, y_o, \{h_j\}_{j=0}^{r-1}) \quad (0 \le r < k) \text{ kezdő értékek,} \\ (1.3) \quad & \sum_{i=0}^{k} \alpha_i y_{n+1}/h_{n+k-1} = \Phi_f\, x_n, \dots, x_{n+k}; y_{n+k}, y_{n+k-1}, \dots, y_n) \\ & 0 \le n \le N-k \end{aligned}$$

ahol $\{\alpha_i\}_{i=0}^{k} \subset R$ adott konstansok és $\Phi \in ([x_o, b]^{k+1} \times R^{m(k+1)}, R^m)$ adott függvény, mely az $y_n, \dots, y_{n+k}$ vektorváltozókban egyenletesen kielégiti a Lipschitz feltételt.

Ekvidisztans esetben (1.3)-at más ekvivalens formákban is használjuk. Például ekkor a jobb oldalon $\Phi_f(x_n, y_{n+k}, y_{n+k-1} \dots, \dots, y_n, h)$-t irunk.

Egylépéses módszerek esetén az irodalomban elterjedtebb az

$$y_{n+} - y_n = h_n \emptyset_f\ (x_n, y_n, y_{n+1}, h_n)$$

formula használata, ahol az előbbi feltételeknek az felel meg, hogy $\Phi_f \in C([x_o, b] \times R^m \times R^m \times [0, h^*], R^m)$ és az y_n, y_{n+1} változók szerint Φ egyenletesen kielégiti a Lipschitz feltételt, ahol $h^* > 0$ konstans. Ha a Φ nem függ az x_{n+k} értéktől, akkor a módszer explicit, különben implicit. Definiáljuk a fenti módszerek (első) karakterisztikus polinomját a következőképpen

$$(1.4) \qquad \rho(x) = \sum_{i=0}^{k} \alpha_i x^i.$$

A továbbiakban az (1.3) módszereket (ρ, Φ_f) alakban is emlitjük majd.

Az (1.3) módszerosztály által tartalmazott ismertebb módszerek: pl.

a/ lineáris k-lépéses módszer (LMN) esetén:

$$\Phi_f = \sum_{i=0}^{k} \beta_i^{(n)} f(x_{n+i}, y_{n+i}) \tag{1.5}$$

Ha a $\beta_i^{(n)}$ konstansok n-től függetlenek, akkor szokásos a második karakterisztikus polinomot is definiálni

$$\sigma(x) = \sum_{i=0}^{k} \beta_i x^i ;$$

b/ prediktor-korrektor módszerek $(P/EC/^{m}E)$ $(m=1)$

$$\Phi_f = \sum_{i=0}^{k-1} \beta_i f(x_{n+i}, y_{n+i}) + \tag{1.6}$$

$$+ \beta_k f\{x_{n+k}, \frac{1}{\alpha_k^*} \sum_{i=0}^{k-1} [-\alpha_i^* y_{n+i} + h \beta_i^* f(x_{n+i}, y_{n+i})]\}$$

ahol α_i^*, β_i^* prediktor formula együtthatói.

c/ Runge-Kutta módszerekre (m-pontos RK)

$k=1$

$$\Phi_f = \sum_{r=1}^{m} c_r k_r \tag{1.7}$$

$$k_r = f\{x_n + h a_r, y_n + h \sum_{s=1}^{m} b_{rs} k_s\} \qquad (1 \le r \le m).$$

Ha a k_r számitásánál az összegzés csak $r-1$-ig megy, a módszer explicit, különben implicit.

Az (1.5) módszerek speciális esetként tartalmazzák

a/ pl. az Adams-tipusu módszereket:

$$y_{n+k} - y_{n+k-1} = h_{n+k-1} \sum_{i=0}^{k} \beta_i^{(n)} f(x_{n+i}, y_{n+i}) \tag{1.8}$$

és

b/ az un. retrográd (BDF) differencia formulákat:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = h_{n+k-1} \beta_k^{(n)} f(x_{n+k}, y_{n+k}). \tag{1.9}$$

Szükségünk lesz a következő fogalmakra.

1. Definició Az $e_n = y_n - y(x_n)$ különbséget az (1.3) osztályba tartozó módszerek globális hibájának nevezzük az $x = x_n$ pontban $(0 \le n \le N)$.

2. Definició Az (1.3) osztályba tartozó k-lépéses módszerek $x_{n+k} \in [x_o, b]$ pontbeli képlethibájának a következőt nevezzük

$$\tau\,(x_n, \{h_{n-j}\}_{j=1}^{k=1}) = y_n - y(x_n), \quad (0 \le n < k,\ k > 1)$$

$$(1.10) \quad \tau\,(x_{n+k-1}, \{h_{n+k-j}\}_{j=1}^{k}) = \sum_{i=0}^{k} \alpha_i\, y(x_{n+i}) - h_{n+k-1}\, \Phi_f(x_n, \ldots, x_{n+k}; y(x_{n+k}), \ldots, y(x_n))$$

$$0 \le n \le N-k$$

ahol $y(x)$ az (1.1), (1.2) probléma elméleti megoldása.

A $k=1$ esetben az x_{n+1} pontbeli képlethibára a $\tau(x_n, \{h_{n+1-j}\}_{j=1}^{1})$ jelölés helyett a hagyományos, ettől csak előjelben különböző

$$\tau(x_n, h_n) = -\tau(x_n, \{h_{n+1-j}\}_{j=1}^{1})$$

definiciót használjuk és x_n pontbeli képlethibáról beszélünk.

Ekvidisztáns lineáris k-lépéses módszerek esetén $(k>1)$ pedig a $\hat{\tau}(x_n, h) = \tau(x_{n+k-1}, \{h_{n+k-j}\}_{j=1}^{k})$ jelölést alkalmazzuk, ahol $h_n = h_{n+1} = \ldots = h_{n+k-1} = h$.

Vizsgáljuk most az alábbi un. perturbált Cauchy problémát

$$(1.11) \quad \begin{aligned} y' &= f(x, y), \\ y_n(x_n) &= y_n \end{aligned}$$

amelynek elméleti megoldását jelölje $y_n(x)$.

3. Definició Az (1.3) alaku módszerek $x_{n+1} \in [x_o, b]$ pontbeli lokális hibájának nevezzük az

$$y_n(x_{n+1}) - y_{n+1} \qquad (0 \le n \le N-1)$$

mennyiséget.

4. Definició Az (1.3) osztályba tartozó közelitő módszer konvergens tetszőleges $\{\Delta_N\} \subset \pi$ felosztássorozatra nézve, ha az (1.1), (1.2) Cauchy-problémára

$$\max_{0\le n<k} \| y_n - y_o \| \to 0 \qquad (\|\Delta_N\| \to 0),$$

(1.12)

$$\max_{0\le n\le N} \| y_n - y(x_n) \| \to 0 \qquad (\|\Delta_N\| \to 0).$$

5. Definició Az 1.3 osztályba tartozó közelitő módszer konzisztens a $\{\Delta_N\} \subset \pi$ felosztássorozaton, ha

$$\max_{0\le n\le N-1} \| \tau(x_n, \{h_{n-j}\}_{j=1}^{k}) \| \to 0 \qquad (\|\Delta_N\| \to 0).$$

$$\sum_{n=0}^{k-1} \| \tau(x_n, \{h_{n-j}\}_{j=1}^{k-1}) \| + \max_{0\le n\le N-k} \| \tau(x_{n+k-1}, \{h_{n+k-j}\}_{j=1}^{k}) \| \to 0$$

$(\|\Delta_N\| \to 0)$, ha $k>1$

illetve

$$\max_{0\le n\le N-1} \| \tau(x_n, h_n) \| \to 0 \qquad (\|\Delta_N\| \to 0) \quad \text{ha} \quad k=1.$$

A konzisztencia p-ed rendü a $\{\Delta_N\} \subset \pi$ felosztás sorozaton, ha p az a maximális pozitiv szám, amelyre a fenti mennyiségek

$$O(\|\Delta_N\|^P)$$

nagyságrendüek.

Az (1.5) módszert formálisan konzisztensnek nevezzük, ha

$$\sum_{i=0}^{k} \alpha_i = 0, \quad \sum_{i=1}^{k} \alpha_i \sum_{j=1}^{i} h_{n+j-1} = h_{n+k-1} \sum_{i=0}^{k} \beta_i^{(n)}$$

teljesül $(0\le n\le N-k)$.

A módszer rendjét a következőképpen definiáljuk. Jelölje $p(f)$ a módszer konzisztencia rendjét az (1.1), (1.2) problémán az ekvidisztáns felosztássorozatokra nézve, és legyen A: az R^{m+1} téren értelmezett analitikus függvények halmaza.

6. Definició A módszer rendjének (globális rend) a

$$p = \min \{p(f) \mid f \in A\}$$

számot nevezzük.

2. A KÖZELITŐ MÓDSZEREK KONVERGENCIÁJA ÉS STABILITÁSA

A továbbiakban először az ekvidisztans felosztássorozatokra kapott konvergencia és stabilitás eredményeket foglaljuk össze.

1. Tétel Ha egy (1.3) osztályba tartozó közelitő módszer konvergens, akkor konzisztens.

(Bizonyitás az LMM esetre Henrici (1962), szélesebb osztályokra ld. Butcher (1966) és Spijker (1966)).

Megjegyezzük, hogy a konzisztencia nem elégséges feltétele az (1.3) tipusu módszerek konvergenciájának. Tekintsük a következő perturbált kezdeti értékproblémát

$$z' = f(x,z)+\delta(x),$$
$$z(x_0) = y_0 + \delta_0, \qquad x \in [x_0, b],$$

ahol $(\delta(x), \delta_0)$ a perturbáció és $z(x)$ a perturbált megoldás.

7. Definició (Hahn (1967)), Stetter (1971)). Legyen $(\delta(x), \delta_0)$, $(\delta^*(x), \delta_0^*)$, tetszőleges két perturbáció és legyen $z(x)$, $z^*(x)$ a kapott perturbált megoldások. Az (1.1), (1.2) kezdeti értékproblémát totálisan stabilisnak nevezzük, ha létezik olyan pozitiv S konstans, hogy minden $x \in [x_0, b]$ esetén teljesül

$$\| z(x) - z^*(x) \| \le S\varepsilon,$$

hacsak $\| \delta(x) - \delta^*(x) \| \le \varepsilon$ és $\| \delta_o - \delta_o^* \| \le \varepsilon$

(Ha az f leképezés Lipschitz feltételnek tesz eleget valamely L konstanssal, ahogy azt feltételeztük, akkor ebből már következik, hogy a DE totálisan stabilis ld. Gear (1971)).

Vizsgáljuk most az (1.3) módszerek alábbi tipusu perturbációit

$$z_r = s_r(f, y_o, \{h_j\}_{j=1}^{r-1}) + \delta_r, \qquad (0 \le r < k)$$

$$\sum_{i=0}^{k} \alpha_i z_{n+i}/h = \Phi_f(x_n, z_{n+k}, \ldots, z_n h) + \delta_{n+k}, \qquad (0 \le n \le N-k)$$

ahol $(\delta_n | n=0,1,\ldots,N)$ a perturbáció és $(z_n | n=0,1,\ldots,N)$ a kapott perturbált megoldás.

8. Definició Legyen $(\delta_n | n=0,1,\ldots,N)$, $(\delta_n^* | n=0,1,\ldots,N)$ tetszőleges két perturbációsorozat és legyenek $(z_n | n=0,1,\ldots,N)$, $(z_n^* | n=0,1,\ldots,N)$ a kapott perturbált megoldások. Ha létezik olyan h_o és s pozitiv konstans, hogy minden $h \in [0, h_o]$ és $\| \delta_n - \delta_n^* \| \le \varepsilon$ $(0 \le n \le N)$ esetén

$$\| z_n - z_n^* \| \le s\varepsilon \quad (0 \le n \le N),$$

akkor az (1.3) módszert <u>D-stabilis</u>nak (vagy zéró stabilisnak) nevezzük (Dahlquist (1956)).

9. Definició Az (1.3) módszer kielégiti a <u>gyökfeltételt</u>, ha a $\rho(\Theta)$ un. karakterisztikus polinom gyökei az egységkörön vagy azon belül helyezkednek el, és a körön elhelyezkedő gyökök egyszeresek.

Ismertek a következő tételek:

2. Tétel Az (1.3) osztályba tartozó közelitő módszer D-stabilis akkor és csak akkor, ha kielégiti a gyökfeltételt.

3. Tétel Az (1.3) osztályba tartozó közelitő módszer akkor és csak akkor konvergens, ha konzisztens és D-stabilis.

4. Tétel Ha az (1.3) alaku módszer konzisztenciája p-ed rendü és a módszer D-stabilis, akkor a konvergencia is p-ed rendü lesz (Hall, Watt (1976)).

Megjegyezzük, hogy az (1.3) osztályba tartozó konzisztens egylépéses módszerek szükségképpen kielégitik a gyökfeltételt, minthogy a $\rho(\theta)$ karakterisztikus polinomnak egyetlen gyöke van a $\theta_1=+1$. Igy igaz a

5. Tétel Az (1.3) osztályba tartozó egylépéses módszer (ahol $k=1$) akkor és csak akkor konvergens, ha konzisztens.

Összefoglalva: a konzisztencia befolyásolja a lokális hiba nagyságát, a D-stabilitás pedig a hibaterjedést, ha $h \to 0$. Számos módszerosztályra bizonyitott tény, hogy

"stabilitás + konzisztencia = konvergencia".

Az LMM esetben a következő eredmény ismert.

6. Tétel A D-stabilis k-lépéses lineáris módszerek maximális rendje $k+1$, ha k páratlan és $k+2$, ha k páros. (Henrici (1962)).

A nem-ekvidisztans esetben (változó lépéshosszra) ismert legáltalánosabb konvergencia és stabilitási tételek a következők.

Az (1.3) alaku egylépéses módszerekre tegyük fel, hogy az (1.1) (1.2) Cauchy probléma megoldásán folytonosak és kielégitik a Lipshitz feltételt az y_n és y_{n+1} vektorváltozók szerint.

7. Tétel Az (1.3) módszercsaládba tartozó egylépéses módszer, melyre teljesülnek a fenti feltételek, akkor és csak akkor konvergens minden $x\in[x_0,b]$ és tetszőleges $\{\Delta N\}_{N=1}^{\infty} \subset \pi$ felosztássorozatra, ha konzisztens, azaz, ha $\Phi_f(x,y,y,0) = f(x,y)$.

(Galántai (1978)).

Az (1.5) alaku változó lépéshosszu lineáris többlépéses módszerekről tegyük fel most, hogy kielégitik a gyökfeltételt; az $\alpha_k=0$ és a $\{\beta_i^{(n)}\}_{i=0}^{k}$ együtthatók pedig a $h_n,\dots,h_{n+k-1}$ lépéshossz homogén függvényei. Minden esetben feltesszük, hogy $y_i \to y_0$ ($\|\Delta_N\| \to 0$, $0 \leq i < k$).

8. Tétel Ha a $\{\Delta_N\}_{N=1}^{\infty} \subset \pi(x \in [x_0, b]$ felosztássorozat olyan, hogy $\sup_N \max_{n,i} |\beta_i^{(n)}| \leq c_1 < +\infty$, akkor az (1.5) formálisan konzisztens módszer konvergens. (Galántai (1978)).

3. A MÓDSZEREK HIBABECSLÉSE

Itt csak rövid összefoglalást kivánunk adni a különböző tipusu hibabecslő eljárásokról. Részletesebben például Hall, Watt (1976) vagy Gear (1971) könyvében találhatók meg.

Megjegyezzük, hogy általában p-ed rendü módszerre elég sima jobb oldalakkal rendelkező DE esetén a lokális hiba a következő alaku

$$\tau_n(x_n, h_n) = \Psi_f(x_n, h_n) h_n^{p+1} + O(h_n^{p+2}) \tag{1.13}$$

ahol Ψ_f az un. <u>főhibatag</u> függvény. A Ψ_f tagot szokás becsülni, lehetséges választások pl. (van der Houwen (1977))

$$\Psi(x_n, h_n) = c_n,$$

$$\Psi(x_n, h_n) = B_n x + c_n,$$

$$\Psi(x, h_n) = A_n x^2 + B_n x + c_n,$$

$$\Psi(x, h_n) = A_n x, \qquad \text{stb.}$$

ahol c_n, B_n, A_n konstansok.

Vizsgáljuk meg röviden a skalár esetben a perturbációk hatását az (1.1), (1.2) problémára. Tegyük fel, hogy $z(x)$ kielégiti az alábbi egyenletet

$$(1.14)\qquad z'(x)-f(x,z(x)) = \theta\delta(x), \qquad (x\in[x_o,b]),$$
$$z(x_o) = y_o+\theta\delta_o \qquad (\theta \text{ kicsi})$$

Legyen (1.15) $z(x)=y(x)+\theta e(x)+O(\theta^2)$ akkor felhasználva Taylor tételét az (1.14) egyenletből kapjuk

$$y'(x)+\theta e'(x)-f(x,y(x))-f_y(x,y(x))\theta e(x)=\theta\delta(x)+O(\theta^2)$$
$$y(x_o)+\theta e(x_o)= y_o+\theta\delta_o+O(\theta^2),$$

ahol $\quad f_y = \{\frac{\partial f}{\partial y}\}$.

Igy az $e(x)$ függvénynek ki kell elégitenie az alábbi lineáris egyenletet:

$$(1.16)\qquad e'(x)-f_y(x,y(x)(e)x) = \delta(x),$$
$$e(x_o) = \delta_o.$$

Azaz ha y az (1.1), (1.2) és a z az (1.14) által definiált, valamint $e(x)$ (1.16)-ból, akkor (1.15) teljesül. Az $e(x)$ megoldás felirható zárt alakban:

$$e(x) = E(x_o x)\delta_o + \int_{x_o}^{x} E(u,x)\delta(u)du,$$

ahol

$$E(u,x) = \exp[\int_u^x f_y(t,y,(t))dt]$$

A fenti képletből könnyen belátható, hogy a $\delta(u)$ perturbáció hatása az u pontban az $E(u,x)$-től függ, amely lehet csökkenő vagy növekvő függvény. A legegyszerübb esetben $y'=\lambda y$, igy $f_y=\lambda$, akkor $E(u,x)=\exp(\lambda(x-u))$. Ha $\lambda>0$, akkor egy u-hoz közeli pontban levő lokális hiba hatása a globális hibára az x pontban exponenciálisan nő, ha x nő. Ha $\lambda<0$, akkor a forditottja áll fenn.

Diszkrét esetben, ha ekvidisztans a lépéshossz, ismert az alábbi egyszerü tétel.

9. Tétel Ha az (1.3) alaku módszer konzisztenciája p-ed rendü és a módszer D-stabilis, akkor a konvergencia is p-ed rendü lesz (Hall, Watt (1976)).

Könnyen belátható, hogy p-ed rendü módszerre

$$y_n = y(x_n)+h^p e(x_n) + O(h^{p+1}) \qquad (0 \le n \le N)$$

ha

$$\tau_n = s_n(f,y_o)-y(x_n) = O(h^{p+1}) \qquad (0 \le n < k)$$

és

$$\tau_{n+k} = h^p \Psi(x_{n+k})+O(h^{p+1}) \qquad (0 \le n \le N-k)$$

ahol

$$e'(x)-f_y(x,y(x))e(x) = -\Psi(x), \qquad e(x_o) = 0.$$

Ha az 1.3 alaku módszer D-stabilis, akkor y_n és $y(x_n) + h^p e(x_n)$ eltérése $O(h^{p+1})$ nagyságrendü.

Ez az eredmény azt jelenti, hogy ha egy módszer lokális hibájának főtagja $h^p\Psi(x_n)$ az x_n pontban, akkor a globális hiba főtagja $h^p e(x_n)$ nagyságu és a pontos megoldástól való eltérés ugyanakkora mint ha $h^p\Psi(x_n)$ taggal perturbáltuk volna az (1.1), (1.2) egyenletet feltéve, hogy a kezdeti feltételek $p+1$ rendig pontosak (általánosabb megfogalmazásban lásd Stetter (1973)).

Néhány ismert módszer leirását megadjuk az 1. táblázatban, amely konkrét hibabecslési eljárásokat is tartalmaz. (1. melléklet)

3.1 Az egylépéses módszerek képlethibájának becslése

A legismertebb hibabecslési módszerek a következő tipusuak: Legyen p a módszer rendje.

a/ <u>Extrapolációs módszer</u>

Legyen $x_{n+1} = x_n+h$,

$$x_{n+2} = x_n+2h \qquad f \in C^{p+1}[x_o,b]$$

Jelölje továbbá $\bar{\bar{y}}_{n+2}$ az x_n ponttól 2h lépéshosszal számitott közelitő megoldást. Ekkor a képlethiba becsléseként

$$\tau(x_{n+2},h) \approx \frac{\bar{\bar{y}}_{n+2} - y_{n+2}}{2^{p+1}-2}$$

(Henrici (1962)) mennyiséget fogadjuk el, amely az $n=0$ esetben $O(h^{p+2})$ nagyságrendü becslését adja a léplethiba fõtagjának.

b/ Beágyazási tipusu módszerek

Minden integrálási lépést kétszer hajtunk végre, elõször egy p-ed rendü (1.7) alaku formulával, majd egy $p+1$-ed rendü képlettel számolva, melyre a k_r értékek ugyanazok, csak uj tagok is jönnek hozzá.

Ily módon a két közelités különbsége segitségével becsülhetõ a p-ed rendü közelités képlethibája.

A legismertebb beágyazási tipusu formulákat pl. England (1967, 1969); Shintani (1965), Fehlberg (1968, 1969); stb. konstruálták.

c/ Többlépéses hibabecslési módszerek

$p+1$-ed rendü többlépéses formulával becsülhetõ a képlethiba az alábbi módon:

$$\tau(x_n,h) \approx \{ \sum_{i=-k+1}^{1} \alpha_i y_{n+i} (h-\beta_i f(y_{n+i}))\}/\rho'(1)$$

(Ceschino, Kuntzman (1963)).

3.2 A k-lépéses prediktor-korrektor módszerek képlethibájának becslése

Milne módszere

Alkalmazható, ha a prediktor és korrektor formula azonos rendü.

A hiba alakja p-ed rendü módszerekre, feltéve, hogy az elõzõ y_r értékek pontosak $(0 \leq r < k)$ az x_{n+k} pontban a prediktorra $c^* h^{p+1} y^{(p+1)}(x_{n+k}) + O(h^{p+2})$, a korrektorra $c h^{p+1} y^{(p+1)}(x_{n+k}) + O(h^{p+2})$,

(C és C^* a formulák együtthatóitól függő konstansok).

Feltesszük, hogy a kezdő értékekben a hiba h^{p+1} nagyságrendü és a lépéshossz konstans, (vagy szakaszosan konstans), valamint a prediktor karakterisztikus polinomja ρ^* és a korrektor karakterisztikus polinomja ρ olyan, hogy $\rho^*(\theta)=0$, ha $\|\theta\| = 1$ és $\rho(\theta)=0$. A következő becslést kapjuk:

$$\hat{\tau}(x_{n,h}) = h^p\Psi(x_{n+k})+O(h^{p+1}) = Ch^p y^{(p+1)}(x_{n+k})+O(h^{p+1}) =$$

$$= K(y_{n+k}^{(PR)} - y_{n+k}^{(C)})/h,$$

ahol

$$K = \frac{a_k^* C/\rho'(1)}{C/\ '(1)-C^*/\rho^{*\prime}(1)} .$$

(Lásd Henrici (1962)).

Gear módszere

Legyen $\underline{a} = [y,hy',\dots,h^p y^{(p)}/p!]^T$,

ha a korrektor p-ed rendü.

A főhibatag p-ed rendü módszere

$$C_{p+1}p!\,\|\nabla a_p/\omega\|_2,$$

alaku, ahol ω sulykomponens, $\|\ \|$ az L_2 norma és ∇a_p az $\underline{a}$ vektor: utolsó komponensének retrográd differenciája. (lásd 1. táblázat).

4. DÖNTÉSI KRITÉRIUM, LOKÁLIS HIBABECSLÉSI ELV

Legyen az (1.1), (1.2) kezdeti érték feladat megoldására az x_n pontban ismert egy $y_n=y(x_n)$ közelités és az $x_{n+1}=x_n+h_{n+1}$ pontban szeretnénk kiszámitani az y_{n+1} értéket. Legyen $y_n(x)$ pontos megoldása az (1.11) Cauchy problémának és ε tetszőleges előre megadott pozitiv hibakorlát. Az (1.1), (1.2) Cauchy-probléma ε pontossággal való numerikus megoldásán azt értjük, hogy a globális hiba $\| e_n \| \leq \varepsilon$. A gyakorlatban ennek biztositására a következő két kritérium egyikét használják egy adott integrálási lépés elfogadásához.

Az un. lépésenkénti hibaellenőrzés (error per step) esetén vizsgáljuk a

$$\| y_n(x_{n+1})-y_{n+1} \| \leq \varepsilon^* \qquad (\varepsilon^* < \varepsilon) \tag{1.17}$$

feltétel teljesülését.

A másik módszer az egységlépésenkénti hibaellenőrzés (error per unit step), ekkor a

$$\| y_n(x_{n+1})-y_{n+1} \| \leq \varepsilon^* h_n \qquad (\varepsilon^* < \varepsilon) \tag{1.18}$$

feltételt kötjük ki a lépés elfogadására.

Az (1.17), (1.18) feltételek ellenőrzését lényegében a következő meggondolások alapján végezzük. Explicit egylépéses módszerek esetén a lokális hiba megegyezik a módszer (1.11) perturbált Cauchy problémára vonatkozó képlethibájával, azaz

$$y_n(x_{n+1})-y_{n+1} = \tau_n(x_n,h), \quad \text{ahol}$$

$$\tau_n(x,h) = y_n(x)+h\emptyset_f(x,y_n(x),h)-y_n(x+h).$$

Az $n=0$ esetet jelölje röviden $\tau(x,h)$. Implicit módszerek esetén belátható, hogy

$$y_n(x_{n+1})-y_n = \tau_n(x_n,h)+O(h^{p+2}).$$

A $\tau_n(x_n,h)$ perturbált képlethiba ("lokális hiba") becslésére általában valamely az előzőekben ismertetett képlethibabecslő eljárást használunk. Ismert eredmény, hogy ha az eredeti (1.1), (1.2) problémára vonatkozó $\tau(x,h)$ képlethiba kielégiti a

$$\| \tau(x,h) \| \le Dh^{p+1} \qquad (x, x+h \in [x_o, b],\ p > 0)$$

feltételt, akkor $\Delta_N \in \pi$, $\| \Delta_N \| \le h^*$ esetén létezik olyan $C>0$ konstans, hogy

$$\| e_n \| \le ch^p.$$

A fenti eredmény alapján alakult ki az a gyakorlat, hogy az n-edik lépésben vett képlethibabecslést vetjük össze az előirt pontossági feltételekkel. $(0 \le n \le N)$.

Valójában azonban nem a $\tau(x_n, h_n)$ képlethibát hanem a már emlitett $\tau_n(x_n, h_n)$ perturbált képlethibát (lokális hibát) becsüljük. Minthogy a $\tau_n(x_n, h_n)$ képlethiba az eredeti $y(x_n)$ megoldástól $O(h^p)$ nagyságrenddel különböző $y_n(x_n) = y_n$ kezdeti feltételhez rendelt Cauchy problémára vonatkozik, elvileg lehetséges, hogy a $O(h^{p+2})$ pontossággal becsült perturbált képlethiba ellenére a globális hiba eggyel kisebbrendü lesz, mint az elméletileg garantált. Ez azonban nincs igy. Ennek a problémának az elvi megalapozása két részből áll; egyrészt a képlethiba és képlethibabecslő eljárás viszonyának vizsgálatából, másrészt az un. lokális hibabecslési elv igazolásából. De ezekkel itt nem foglalkozunk behatóbban (ld. pl. Galántai (1978)).

10. Tétel ha $\Delta_N \in \pi$ $(\| \Delta_N \| \le h^*)$ olyan, hogy az (1.3) alaku egylépéses módszer $(k=1)$ hibájára

$$(1.19) \qquad \| \tau_n(x_n, h_n) \| \le h_n \varepsilon^* \quad (\| \tau_n(x_n, h_n) \| \le \varepsilon^*)$$

fennáll, akkor létezik $c=c(f,\emptyset)>0$ konstans $[c=c(N)]$ szám, amelyre

$$(1.20) \qquad \| y_n - y(x_n) \| \le c\varepsilon^* \qquad (0 \le n \le N).$$

(Galántai (1978)).

Ez a tétel indokolja az un. lokális hibabecslési elvet, azaz, hogy a $\| e_n \| < \varepsilon$ feltétel teljesülése helyett az (1.17) vagy (1.18) feltétel teljesülését vizsgáljuk, amely a lokális hibára vonatkozik.

5. LÉPÉSHOSSZVÁLTÁS

Egy adott integrálási lépést sikeresnek fogadnak el, ha teljesül rá az (1.17) vagy (1.18) feltétel a megadott ε pontossággal. Jelölje p-ed rendü módszerre az $\| y_n(x_{n+1})-y_{n+1} \|$ eltérésre adott becslést "est". Ha a lépés nem elfogadható pontosságu, akkor csökkentenünk kell a lépéshosszt általában ugy, hogy a feltétel teljesüljön, de a lépéshossz a lehető legnagyobb legyen. Az

$$(1.21) \qquad |est| = |h_n^{p+1}\Psi(x_n,y_n)|$$

formulából látható, hogy a lépést a következő h_n', uj lépéshosszal kell megismételnünk:

$$h_n' = |\varepsilon/est|^{\frac{1}{p+1}} h_n,$$

mert

$$|h_n'^{\,p+1}\Psi(x_n,y_n)| = \frac{\varepsilon}{|est|} h_n^{p+1} |\Psi(x_n,y_n)| = \varepsilon .$$

Ha a lépést elfogadtuk, akkor meg akarjuk határozni azt a legnagyobb lépéshosszt, amellyel sikeresen számolhatunk a következő lépésben.

A következő megfigyelésből

$$y_{n+1}(x_{n+2})-y_{n+2} = h_{n+1}^{p+1}\Psi(x_{n+1},y_{n+1})+O(h_{n+1}^{p+2}) =$$

$$= h_{n+1}^{p+1}\Psi(x_n,y_n)+O\ h_{n+1}^{p+2} ,$$

látjuk, hogy a megfelelő érték

$$h_{n+1} = |\varepsilon/est|^{\frac{1}{p+1}} h_n .$$

Igy ugyanaz a technika használható sikeres és sikertelen lépés esetén is az uj lépéshossz kiválasztására. Az ilyen módon meghatározott mennyiségeket "lokálisan optimális" lépéshossznak fogjuk nevezni. (lásd pl. Jackson (1978)).

Óvatosnak kell azonban lennünk a lokálisan optimális lépéshossz gyakorlati felhasználásakor. Az (1.21) közelitő egyenlőség nagyon pontatlan lehet, ha pl. a lokális hibabecslés pontatlan, vagy mert az (1.13) h-szerinti sorfejtésének z-nél magasabb rendü tagjai nem elhanyagolhatók. A tapasztalatok szerint ez különösen sikertelen lépés esetén állhat fenn. Ha a DE nem elég sima, akkor (1.13) nem igaz, a h_n hatványai redukálódnak annak megfelelően, mennyire sima a DE megoldása. Ha a megoldás sima, még akkor is előfordulhat, hogy a főhibatag $\Psi(x,y)$ elég gyorsan változik, és a fent leirt módon felhasználva, egy sikeres lépés után gyenge közelitést kaphatunk. Általában a lokálisan optimális lépéshossz egy tört-részét használják a gyakorlatban. Pl. az irodalomban DIFSUB név alatt ismert programban (Gear (1971)) a lokálisan optimális lépéshossz 0.99-ed részét használják; de általában ez 0.5; vagy 0.8 értéktől 0.9-ig változik.

Megjegyezzük még, hogy a globális hiba nem reguláris viselkedésü, ha nem a közelitőleg optimális lépéshosszt használjuk. Pl. Ebben az esetben előfordul, hogy a hibakorlát csökkentésével a globális hiba kicsit változik, ha egyáltalán változik.

Példa: az irodalomból RKGS (CDC programkönyvtár) néven ismert Runge-Kutta-Gill módszer programját alkalmazzuk a következő DE-re

$$\begin{aligned} y_1' &= y_2, & y_1(0) &= 0, \\ y_2' &= -y_1, & y_2(0) &= 1, \end{aligned}$$

a $[0, 8\pi]$ intervallumon. Az eljárás lépéshosszfelezéssel ill. duplázással változtatja a lépéshosszt. Bemutatjuk a 2. táblázatban a megoldás két komponensének hibáját a 8π pontban különböző ε hibakorlátok esetén.

$-\log_{10}$	y_1 hibája	y_2 hibája	függvénybehivások száma
1	-8.2(-4)	-1.7(-4)	562
2	-8.2(-4)	-1.7(-4)	562
3	-8.2(-4)	-1.7(-4)	562
4	-8.2(-4)	-1.7(-4)	562
5	-6.6(-5)	-8.7(-6)	1108
6	-5.0(-5)	-5.4(-6)	1119
7	-1.6(-6)	-1.3(-7)	2226
8	1.4(-6)	1.5(-8)	4411
9	1.5(-6)	1.7(-8)	4433
10	1.7(-6)	1.1(-8)	8840
11	1.7(-6)	5.5(-9)	17559
12	1.7(-6)	5.5(-9)	17647

2. táblázat Példa

Az alábbiakban Gear, Tu (1974) Gear, Watanabe (1974) eredményeit összegezzük az (1.3) módszerosztályra ill. annak egyes speciális módszereire vonatkozóan.

10. Definició <u>Lépéshosszválasztási sémának</u> nevezzük az olyan függvényt, amelyre

$$h_n = h\theta(x_n, h),$$

ahol tetszőleges $h>0$, $0\leq x\leq b$; $1\geq\theta(x,h)\geq C>0$. A legismertebb a következő két lépéshosszváltoztatási technika.

Az un. <u>interpolációs technika</u> a h_n lépéshosszal való számoláskor az uj lépéshosszra ugyanolyan rendü formula beinterpolálásából áll, azaz a/ interpolálás az ismert közelitésekből, hogy az $\tilde{x}_{n-i} = x_n - ih_n$ $(k > i \geq 0)$ pontokban megkapjuk a megoldás közelitését.

b/ ezeken az $\tilde{x}_{n-i}$ pontokon alapuló fix-lépéses formula segitségével megkeressük a közelitő megoldást az $x_{n+1} = x_n + h_n$ pontban.

A <u>változó lépéshosszu</u> technika alkalmazásakor k darab nem egyenletesen elhelyezett $(k > i \geq 0)$ pontból kiindulva kiszámitjuk az uj lépést és a k-lépéses formula együtthatóit, hogy az megfelelő rendü legyen.

Ismert (Dahlquist (1956)), hogy a k-lépéses módszer rendje nem érheti el a $2[k/2]+2$ értéket, ha azt akarjuk, hogy a módszer rögzitett lépéshossz esetén stabilis legyen; ezért néhány $\alpha_{i,n}$ és $\beta_{i,n}$ együttható értékét előirjuk. Például, ha azt adjuk meg, hogy $\alpha_{j,n} = 0$ $(j \geq 2)$ esetén és a módszer $(k+1)$-ed rendü legyen, akkor a változó lépéshosszu Adams-Moulton módszer lesz a megfelelő rendü formula.

Ismertek a következő tételek:

11. Tétel A változó lépéshosszu (1.5) módszerek közé tartozó Adams-Basforth (AB), A-B-Moulton (ABM) prediktor-korrektor és A-M módszerek stabilisak* és konvergensek tetszőleges lépéshosszválasztási sémára.

12. Tétel Ha legalább k konstans lépést teszünk a lépéshossz változatások között, akkor a k-lépéses interpolációs AB, ABM és AM módszerek stabilisak és konvergensek.

13. Tétel Ha egy módszer fix-lépéshosszra kielégiti a stabilitási feltételt, akkor tetszőleges olyan lépéshossz választá-

---

* A stabilitás a 11. 12. 13. tételben kicsit eltérő az ismertetett definiciótól, de az f függvényre vonatkozó alkalmas feltételek mellett a két definició ekvivalens.

si sémára stabilis (változó lépéshosszu ill. interpolációs sémára), amely csak kis változást eredményez, azaz amelyre

$$\frac{h_{n+1}}{h_n} = 1+O(h) \qquad (h \to 0).$$

Az interpolációs technika előnye a rugalmas használhatóságában rejlik, (az áttérés az egyik lépéshosszról a másikra általában egy mátrix-vektor szorzás segitségével megoldható (Nordsieck (1962)). De ez a technika nem megfelelő ha magasabb rendü formulákat használunk (pl. Krogh (1973)).

A változó lépéshosszu technika esetén nehézkesebb a lépéshosszváltoztatás utáni számolás.

11. Definició <u>Formula választási</u> sémának nevezzük azt az $I(h,x)$ függvényt, amely definiálja az $[x_n, x_{n+1}]$ intervallumon használt $F_{I(h,x_n)}$ formulát, ahol h a lépéshosszválasztási séma által definiált paraméter.

12. Definició <u>Változó formulát használó módszer</u> a következőkből áll: az $\{F_i\}$ többlépéses formulák halmazából, valamint a $\{\theta_{ji}\}$ formula változtatási operátorok halmazából, a lépéshosszváltoztatási technikából, a lépéshossz- és formulaválasztási sémából.

14. Tétel A változó lépéshosszu technikát felhasználó változó rendü (1.10) Adams módszer stabilis tetszőleges formula választási séma esetén.

6. EFFEKTIVITÁS, A MÓDSZEREK ALKALMAZÁSA

Végül szükséges megemlitenünk a következő az irodalomban eddig kevéssé vizsgált problémát. A módszerek konvergenciájára vonatkozó tételek egy fix módszer adott felosztássorozaton való konveregenciáját biztositják adott $\|\Delta_N\| \to 0$ felosztássorozaton. Jelölje $M = \langle(\rho,\emptyset), B,D,H\rangle$ azokat az eljárásokat, ahol $(\rho,\emptyset)$ az adott (1.3) alaku módszer, B a képlethiba valamely becslése, D adott döntési kritérium és H adott lépéshosszválasztási stratégia. Tekintsük az (1.1), (1.2) feladat $x \in [x_o, b]$ pontbeli megoldását az (1.17) feltétel mellett, azaz a

$$P(\varepsilon) = \left\langle f, x_o, y_o, x; \begin{array}{l} \| \tau_n(x_n, h_n) \| \leq \varepsilon, \ x_{n+1} \in \Delta_N \in \pi \\ \| B(x_n, h_n) \| \leq \varepsilon, \ f \in F \end{array} \right\rangle$$

problémaosztályt (ahol F adott függvényosztály).

13. Definició A $\langle(\rho,\emptyset)$ B,D,H$\rangle$ eljárást <u>effektivnek</u> nevezzük a $P(\varepsilon)$ problémaosztályra nézve, ha vele az osztálynak minden eleme megoldható (Hull (1969)).

Ha a $(\rho,\emptyset)$ módszer konvergens bizonyos tipusu $\{\Delta_N\} \subset \pi$ sorozatokra, akkor az eljárás effektivitása olyan B,D,H kritériumokat követel meg, amely ilyen tipusu $\{\Delta_{N(\varepsilon)}\} \subset \pi$ $(\varepsilon \to 0)$ felosztássorozatot állit elő és

$$(1.22) \qquad \| y_{N(\varepsilon)} - y(x_N) \| \leq \varepsilon \qquad (0<\varepsilon<\varepsilon_o).$$

Az effektivitás vizsgálata azért lényeges, mert a $(\rho,\emptyset)$ módszer konvergenciájából még nem következik, hogy adott $p(\varepsilon) \in P(\varepsilon)$ probléma esetén a $\langle(\rho,\emptyset)$,B,D,H$\rangle$ módszer által számitott közelités globális hibájára teljesül az (1.22) feltétel.

Effektivitási tételeket egyszerübb esetekben ekvidisztans felosztásokon Hull (1969) (lásd alább) és Sacks-Davis (1977) igazoltak. Megjegyezzük, hogy a 10. tétel is interpretálható általános effektivitási tételként.

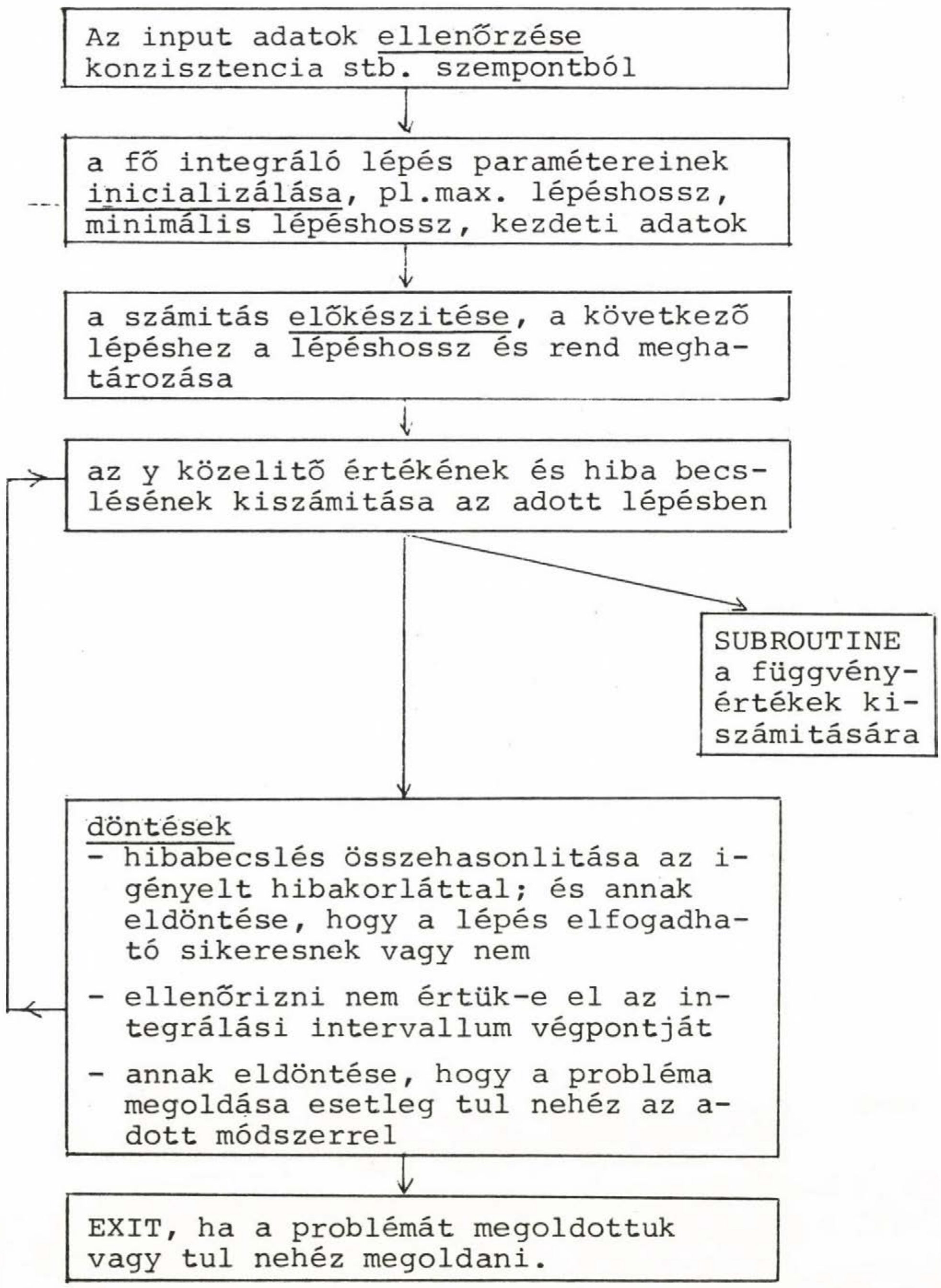
15. Tétel (Hull (1969)). Tekintsük az $y' = Ax, y(x_o) = y_o$ feladatot, (y_o, x_o adott), A mxm dimenziós négyzetes mátrix. Legyen továbbá <u>b</u> adott végpont és egy lépés elfogadhatóságának kritériuma az, hogy a lokális hiba egy adott ε korlát alatt van. Tegyük fel, hogy explicit negyedrendü Runge-Kutta módszert alkalmazunk lépésfelezéses hibabecsléssel.

Akkor a módszer effektiv a fenti problémaosztályon, ha

1/ egylépés eredményét csak akkor fogadjuk el, ha a lokális hibára kapott becslés nem éri el a $3/4h\varepsilon$ értéket és

2/ a h lépéshossz kiválasztási stratégiája olyan, hogy $h \leq h_{max}$, ahol $h_{max} \leq 1/4 \| A \|$ és a stratégia olyan, hogy a módszer segitségével a b végpont elérhető legyen.

Az 1. ábrán bemutatjuk egy DE közelitő megoldásának gyakorlati folyamatát, amelynek segitségével például egy módszer számitógépre vihető.



1. ábra

Az implementálás folyamata

II. MÓDSZEREK ÖSSZEHASONLITÓ VIZSGÁLATA

I. ÁLTALÁNOS SZEMPONTOK

Számos módszer létezik az (1.1) - (1.2) Cauchy-féle kezdeti érték probléma közelitő megoldására, de nem létezik olyan optimális módszer, amely tetszőleges differenciálegyenlet esetén mindig a legjobb eredményt adja.

A különböző szempontokból optimális módszerek kiválasztására számos technika ismeretes. Ezek lényegét világitjuk meg az alábbiakban.

Általában ha numerikus módszereket akarunk összehasonlitani, definiálnunk kell a következőket:

1. A megoldandó problémák osztálya (Jelölje ezt P, elemeit $p \in P$)
2. A módszerek osztálya: M $(m \in M)$
3. Összehasonlitási kritérium: A (p,m) módszer és problémapárnak megfeleltetünk egy $C(p,m)>0$ értéket, amely a p probléma m módszerrel történő megoldása jóságának mértéke.

Ha ezeket definiáltuk, akkor világos, hogy mit értünk azon, hogy az egyik módszer jobb mint a másik. Azt mondjuk, hogy az m módszer jobb, mint az m' módszer a P problémaosztályra a kritériumnak megfelelően, ha $C(p,m) < C(p,m')$.

Azt is mondhatjuk, hogy m a legjobb módszere az M módszerosztálynak a C kritériumra nézve, ha a P problémaosztályra alkalmazva $C(P,m) \leq C(P,m')$; $\forall m' \in M$ esetén.

A problémák osztályát esetünkben az (1.1), (1.2) Cauchy-feladatok alkotják. Ezek leirhatók a következőképpen:

$$p = \langle f, x_0, y_0, b, \varepsilon \ \ldots \ \text{stb.} \rangle$$

ahol f, x_0, y_0 definiálja a problémát matematikailag, ε a lokális hiba felső korlátja és b a végpont, amelyben a megoldást keressük.

A közelitő módszereket a korábbiaknak megfelelően definiálhatjuk (algoritmus, hibabecslés, döntés, lépéshosszválasztás).

Az összehasonlitási kritérium lehet az algoritmussal való számoláshoz és a hibabecsléshez szükséges függvénybehelyettesitések száma.

2. KISÉRLETI VIZSGÁLATOK

Hull és társai (1972) dolgozták ki az első komolyan megalapozott kisérleti (számitógépes) összehasonlitó vizsgálatokat.

Vizsgálataikban ε az egységnyi lépésre megkövetelt lokális hibakorlát. A probléma további jellemzésére bevezetik a h_{max} maximálisan megengedett lépéshosszt. A problémák nem tartalmaznak sem szakadásos jobboldallal rendelkező egyenletet, sem stiff tipusu vagy más, az átlagnál rosszabb egyenletet. Az egyes problémaosztályok a következők

a/ az egyenlet,

b/ kisméretü egyenletrendszer

c/ közepes méretü egyenletrendszer

d/ orbitális egyenletek

e/ magasabbrendü egyenletek.

Minden osztály öt feladatból áll, amelyek mindegyikét vizsgálják 10^{-3}, 10^{-6}, 10^{-9} nagyságu ε hibakorlát esetén. A felhasznált algoritmusok a következők: két Runge-Kutta módszer (Butcher nyolcadrendü módszere (1965) és Shanks módszere (1966)), az Adams-módszer (Gear féle változat (1971), lásd még az 1. táblázatot), Bulirsch, Stoer (1966) extrapolációs módszere. Összehasonlitási kritériumnak a felhasznált függvénybehelyettesitések számához hozzáveszik az un. "overhead" időt, amely alatt a gépidő azon részét értik, amely fennmarad, ha kivonjuk a függvénybehelyettesitésekre elhasznált gépidőt.

Egy statisztikai program segitségével gyüjtik össze ezeket az adatokat, valamint bizonyos megbizhatósági vizsgálathoz szükséges számolásokat (a program kiszámitja a feladat pontos megoldását is és az attól való eltérést).

Az alábbi következtetésekre jutnak:

a/ A változó rendü lineáris többlépéses módszerek általában jobbak, mint a rögzitett rendüek. Általános esetben a leghatékonyabban használhatók, ezért ezek a legelterjedtebbek ma.

b/ A Bulirsch-Stoer módszer kevesebb "overhead" időt igényel, mint az Adams módszerek (azaz, ha a függvénybehelyettesitések viszonylag költségesek kb. $\geq$ 25 aritmetikai müveletet igényelnek komponensenként), akkor az Adams módszer használata célszerübb.

c/ A Runge-Kutta módszerek általában nem versenyképesek a többi vizsgált módszerrel. Kivétel az az eset, ha nincs szükségünk nagy pontosságu megoldásra; akkor ui. az alacsonyabbrendü Runge-Kutta módszerket érdemes használni, ha a függvényértékek számitása nem tulságosan bonyolult.

A továbbiakban egylépéses módszerek összehasonlitásával foglalkozunk.

Shampine és Watts (1976) széleskörü szempontok figyelembevételével elvégzett kisérleti (számitógépes) módszerekkel hasonlitják össze a Runge-Kutta formulák alkalmazását egy adott feladatosztályon.

(2.1) $$y'(x) = \sum_{i,j=0}^{\infty} a_{ij}\, x^i y^j, \quad \text{(konvergens hatványsor)}$$

$$y(0) = 0$$

ahol az a_{ij} értékek véletlen számok, amelyek a [-I, I] szakaszon egyenletes eloszlásuak (mintegy 500 egyenletet generálnak a teszteléshez).

Az (1.7) alaku m-pontos p-ed rendü ($m \leq p \leq 5$) explicit Runge-Kutta módszerek osztályából vizsgálja az alábbiakat:

a/ a klasszikus negyedrendü módszert lépésfelezéssel

b/ Gill (1951) módszerét lépésfelezéssel

c/ England (1969) módszerét, amelynek hibabecslése beágyazási tipusu

d/ Merson negyedrendü módszerét (amely csak lineáris egyenletekre ad aszimptotikusan pontos hibabecslést) harmadrendü módszerként interpretálva, amely lokális extrapolációt használ (igy aszimptotikusan pontos hibabecslés adható)

e/ Zonneveld (1964) ötödrendü módszerét, amelyet lokális extrapolációt használó negyedrendü módszerként interpretálnak

f/ Shintani (1966) negyed-ötödrendü beágyazási tipusu módszerét

g/ Fehlberg (1969) negyed-ötödrendü beágyazási tipusu módszerét két változatban, amelyben csak a módszer paraméterei térnek el egymástól.

Összehasonlitásokat végeznek az adott problémaosztályon, amelyek alapja a pontosság, a hibabecslés minősége, a stabilitás és az általános effektivitás (hatékonyság).

Az adott módszerek <u>pontosságát</u> a főhibatag szerinti összehasonlitással vizsgálják (három különböző normában).

Ahhoz, hogy értelmes összehasonlitást lehessen tenni különböző számu pontból álló formulák esetén, szükségünk van arra, hogy skálázzuk a lépéshosszakat, igy véve figyelembe a nem egyenlő mennyiségü munkát. Igy tekintjük a $h^*=mh$ lépéshosszt, azaz m-pontos módszer esetén a h^q-ad rendü tagokhoz tartozó hibaegyütthatókat beszorozzuk az m^q mennyiséggel. Ez utóbbit fogjuk "skálázott"-nak nevezni a továbbiakban. A 3. táblázatban a hibakonstansokat mutatjuk be "skálázott" és nem skálázott esetben.

A hibabecslés minőségének vizsgálatára végeznek aszimptotikus és nemaszimptotikus összehasonlitásokat. Felteszik, hogy a módszer lokális hibája és annak becslése az alábbi alaku:

$$le = ah^{p+1}+bh^{p+2}+O(h^{p+3}) \qquad \text{(a lokális hiba)}$$

$$est = \alpha h^{q+1}+\beta h^{q+2}+O(h^{q+3}) \qquad \text{(a lokális hiba becslése)}$$

Aszimptotikusan pontos becslés esetén $q=p$ és $\alpha=a$. Megkapják közelitőleg a

$$(2.2) \qquad D(f):=\lim_{h\to 0}\frac{le-est}{h^{p+2}}=b-\beta$$

aszimptotikusan pontos becslést ugy, hogy a lokális hibát és becslését a $h=2^{-N}$ $(N=3,4\ldots)$ sorozatra számitják ki a vizsgált feladatosztályon. Ezt addig végzik, amig a $D(f)$ szukcessziv közelitései egy előirt mennyiségnél kevésbé térnek el. Végül a $|D(f)|$ értékek átlagát és maximális értékét számitják ki minden vizsgált módszerre a (2.1) alaku véletlen módon előállitott feladatokra. A $D(f)$ mennyiség numerikus konvergenciájának elérése után a hibabecslés egy másik mértékét, a

$$(2.3) \qquad R(f):=\frac{D(f)}{\beta} \qquad \text{mennyiséget}$$

számitják ki közelitőleg.

A 4. táblázatban bemutatjuk a $|D(f)|$ és $|R(f)|$ közelitő értékek átlagát és maximumát.

Az 5. táblázatban a felhasznált módszerek rangsorolását próbálják megadni a hibabecslések szempontjából. Megjegyzik, hogy bizonyos esetekben nem lehet egyértelmüen kiválasztani az eljárások közül a legjobbat, legalábbis az extrapoláció nélküli esetben mindegyik hibabecslés megfelelő.

3. táblázat

A pontosság mértéke, a hibatag együtthatói

Módszerek	Skálázatlan				Skálázott		
	p	h^{p+1}	h^{p+2}	h^{p+3}	h^{p+1}	h^{p+2}	h^{p+3}
Merson	3	.010	.021	.028	6.4	6.4(+1)	4.4(+2)
Klasszikus	4	.0022	.0035	.0054	3.5(+2)*	6.3(+3)	1.1(+5)
Gill	4	.0019	.0031	.0048	3.1(+2)	5.5(+3)	9.4(+4)
Zonneveld	4	.20	.51	.70	3.4(+3)	6.0(+4)	6.5(+5)
England	4	.0019	.0031	.0048	1.1(+2)	1.6(+3)	2.3(+4)
Shintani	4	.00089	.0039	.0087	1.5(+1)	4.6(+2)	7.1(+3)
Fehlberg(1)	-	.0061	.015	.025	4.8(+1)	7.0(+1)	7.0(+2)
Fehlberg(2)	-	.0033	.018	.039	2.5(+1)	8.3(+2)	1.1(+4)

4. táblázat

A hibabecslések aszimptotikus összehasonlítása

Módszerek	D(f) átlaga skálázott	D(f) max. skálázott	R(f) átlaga	R(f) max.
Shintani	77.	4.1(+2)	5.4	166.
Fehlberg(2)	102.	4.1(+2)	2.4	85.
Merson	20.	6.9(+2)	1.1	37.
England	78.	7.0(+2)	.72	33.
Fehlberg(1)	47.	2.1(+2)	.69	15.
Gill	202.	1.4(+3)	.43	19.
Klasszikus	271.	1.8(+3)	.39	11.
Zonneveld	151.	9.8(+2)	.036	2.8

---

* a zárójelben levő szám a 10 megfelelő hatvánnyal történő szorzást jelöli.

5. táblázat

A hibabecslések rangsorolása legkevésbé hatékonytól a legjobbig

Nem extrapolációs formulákra	Extrapolációs formulákra
Shintani, Fehlberg(2)	Zonneveld
Merson, Fehlberg(1)	Fehlberg(1)
England, Kalsszikus, Gill	Shintani
Zonneveld	Fehlberg(2)

Shampine és Watts (1976) ismertetett munkájukban effektivitási kérdéseket is érintenek. Összehasonlitják a módszerek hatékonyságát, amikor azok már összemérhető pontosságot értek el. Az egyszerüség kedvéért a magasabbrendü tagokat elhanyagolják a hiba sorfejtésben, igy csak a következő tagot veszik:

$$\text{error} = \varepsilon h^{p+1}$$

Felteszik, hogy a k-adik módszer hibatagja $\varepsilon_k h^{p+1}$, ahol az ε_k értékeket a 3. táblázat tartalmazza. A k-adik módszer h_k lépéshossza, amely mellett a hibatag ugyanaz lesz

$$h_k = \left(\frac{\varepsilon}{\varepsilon_k}\right)^{\frac{1}{p+1}} h$$

Ha N darab h hosszuságu lépésre van szükség, egy adott pont eléréséhez, akkor

$$N_k = \left(\frac{\varepsilon_k}{\varepsilon}\right)^{\frac{1}{p+1}} N$$

darab h_k hosszuságu lépés szükséges ugyanazon pont eléréséhez. A 7. táblázat az egyes módszerek teljes költségét mutatja a függvénybehelyettesitések számával és az un. "overhead" idővel mérve.

7. táblázat

A módszerek költsége

	Függvénybehelyettesítések száma / lépés
Merson	5
Fehlberg	6
Zonneveld, Shintani	7
England	9
Klasszikus, Gill	11

	müveletek száma / lépés	müveletek száma / pont
Klasszikus	31	2.8
Gill	37	3.4
Merson	19	3.8
England	35	3.9
Zonneveld	33	4.7
Fehlberg	30	5.0
Shintani	37	5.3

A különböző szempontok alapján elvégzett vizsgálatok azt mutatják, hogy nem létezik minden kategóriában legjobb módszer.

A fenti vizsgálatok eredményei alapján a szerzők a Fehlberg módszert találták a legjobban és a legáltalánosabban alkalmazhatónak.

3. ELMÉLETI ÖSSZEHASONLITÁSI VIZSGÁLAT

Elméleti módszerekkel vizsgálja Jackson (1978) a Runge-Kutta formulák költségének nagyságát bizonyos feladatosztályon.

Az (1.7) alaku m-pontos p-ed rendü explicit Runge-Kutta módszerek osztályát tekinti a lineáris nem stiff homogén konstans együtthatós problémákra. A lokális hiba becslésére felhasználja a korábban emlitett módszerek közül a lépésfelezési technikát, Ceschino-Kuntzman többlépéses hibabecslését és néhány beágyazási tipusu módszert. Az adott integrálási lépés elfogadhatóságának kritériuma az un. retrográd hibaanalizis elvén alapszik. Adott ε hibakorlátra megköveteli, hogy az eredeti (1.1), (1.2) probléma közelitő megoldása az $[x_0,b]$ integrálási intervallumon a perturbált probléma (mely az eredeti feladat ε-nal korlátos perturbációja) pontos megoldása legyen. Az integrálási lépés elfogadására használt stratégiák közül számos követeli meg az alábbi egyenlőtlenség teljesülését.

(2.4) $\quad a(h)\|E(h)\| \leq \varepsilon$

ahol $E(h)$ a lokális hiba becslése egy adott lépésben, $a(h)$ pedig tetszőleges függvény (sok esetben 1 vagy h).

Modellmódszerünk olyan, hogy az $a(h)$ függvényt helyettesitik egy $b(h)$ un. effektivitási függvénnyel.

Egy problémaosztályra vonatkozó effektivitási függvénynek $b(h)$ nevezik azt a legkisebb (pontszerü) függvényt, amely garantálja, hogy minden $p\in P$ esetén, ha a (2.4) egyenlőtlenség az integrálás minden lépésénél teljesül, akkor a módszer által generált approximáció kielégiti a problémához tartozó elfogadhatósági kritériumot.

A lépéshossz kiválasztásának stratégiája olyan, hogy kiválasztja a (2.4) feltételt kielégitő legnagyobb lépéshosszt, azaz $H(\varepsilon)$ a következő egyenlet egyértelmü gyöke (ld. részletesen később)

(2.5) $\quad b(h) \max\limits_{\|A\| \, \|y_0\| \leq 1} \|E(h)\| = \varepsilon.$

Az adott módszer költségét a következőképpen definiálja

(2.6) $\displaystyle c(\varepsilon) = \max_{\substack{\|A\| \le 1 \\ \|y_0\| \le 1}}$ {az $y' = Ay, y(0) = y_0$ Cauchy probléma megoldásához egy egységnyi lépésre felhasznált azon függvényszámolások minimuma, amelyekre szükség van adott ε hibakorlát mellett}

A lineáris konstans együtthatós homogén problémák (ahol $\|A\| \le 1$ és $\|y_0\| \le 1$) osztályán egy lépés költsége

(2.7) $c(\varepsilon) = F/H(\varepsilon)$

ahol F a formula és a hozzá tartozó hibabecslés által lépésenként elhasznált függvénybehelyettesitések száma.

Szigoru hibakorlátokra a költséggörbék közel lineárisak

$$\log c(\varepsilon) \approx K - \frac{1}{p} \log\varepsilon,$$

ahol K egy konstans, p a módszer rendje.

A $b(h)$ effektivitási függvény közelitését is sikerült előállitaniuk három lépésben. Először összekapcsolják az $[x_{k-1}, x_k]$ intervallumra vonatkozó retrográd hibát és az ugyanerre az intervallumra vonatkozó egységnyi lépéshez tartozó τ_k hibát egy $z(x)$ függvény segitségével. Az egységnyi lépésre vonatkozó hiba a k-adik lépésben

(2.8) $\tau_k(h_k) = [y_k(x_k) - y_k]/h_k,$

ahol $y_k(x)$ az $y'(x) = f(x, y(x))$, $y_k(x_{k-1}) = y_{k-1}$ "lokális" kezdeti érték probléma megoldása.

A $z(x)$ függvényre a következő követelményeket irják elő:

a/ $z(x)$ legyen folytonos az $[x_0, b]$ intervallumon;

b/ $z(x_k) = y_k$ $(k=0,1,\dots n)$, ahol $y_n = \tilde{y}_n$ és

c/ $z'(x) - Az(x) = u_k(x)$, $x \in [x_k, x_{k+1}]$ $(k=1,2,\dots,n)$

ahol $u_k(x)$ az ismeretlen. A szabályozás elméletből ismeretes, (lásd pl. Pontrjagin (1968), hogy tetszőleges olyan $u_k(x)$ függvény, amely kielégiti a fenti feltételt, és minimalizálja a $\max \| u_k(x) \|$, $x \in [x_{k-1}, x_k]$ mennyiséget, az $[x_{k-1}, x_k]$ szakaszon konstans normáju, ezért $u_k(x)$ helyett itt vehető egy u_k konstans vektor (lásd pl. Sedgwick (1973)). Igy $x \in [x_{k-1}, x_k]$ esetén

$$(2.9) \qquad z(x) = e^{(x-x_{k-1})A} y_{k-1} + (x-x_{k-1}) e_1((x-x_{k-1})A) u_k$$

ahol

$$e_1(x) = \begin{cases} \frac{e^x - 1}{x}, & \text{ha} \quad x \neq 0, \\ 1, & \text{ha} \quad x = 0. \end{cases}$$

Az $e_1(x)$ függvény kiterjeszthető négyzetes mátrix függvényekre, létezik a reciproka és analitikus mindenütt a C téren kivéve az $e_1(x)$ gyökeinél $\{2\pi i k;\ k = \pm 1, \pm 2, \ldots\}$. Igy az $e_1^{-1}(x)$ szintén kiterjeszthető mátrixfüggvényekre, feltéve, hogy a mátrixargumentum egyetlen sajátértéke sem gyöke az $e_1(x)$-nek. $e_1^{-1}(hA)$ az $e_1(hA)$ mátrix inverze, ha $\| hA \| \leq 2\pi$. A (2.4) formulába $x = x_k$ értéket helyettesitve, valamint felhasználva, hogy $h_k = x_k - x_{k-1}$ és $\tau_k = (e^{h_k A} y_{k-1} - y_k)/h_k$, kapjuk

$$(2.10) \qquad u_k = -e_1^{-1}(h_k A)\tau_k .$$

A p-ed rendü explicit Runge-Kutta formula esetén τ_k kielégiti a következő egyenletet

$$\tau_k = s(h_k A) \frac{(h_k A)^{p+1}}{h_k} y_{k-1},$$

ahol $s(x)$ egy végtelen konvergencia sugaru hatványsor. A formulához tartozó hibabecslés kielégiti a következő egyenletet

$$E_k = R(h_k A) \frac{(h_k A)^{p+1}}{h_k} y_{k-1},$$

ahol $R(\alpha)$ egy polinom. Ha $R(h_k A)$ nem szinguláris, akkor E_k és

τ_k összekapcsolható a következő egyenlettel:

(2.11) $\tau_k = S(h_k A) R^{-1}(h_k A) E_k$

Felteszik, hogy az $S(x)$ és $R(x)$ polinomoknak nincs közös gyökük.

A (2.10) és (2.11) egyenletből következik, hogy u_k és E_k közelitésére az alábbit kapták:

$$(2.12)\quad \begin{aligned} & b(h) = \max_{\|A\| \le 1} \| e_1^{-1}(hA) S(hA) R^{-1}(hA) \| , \\ & 0 \le h < \min(h_{max}, 2\pi). \end{aligned}$$

A $b(h)$ kiszámitásához $R^{-1}(x)$-et h_{max} konvergenciasugaru hatványsorba fejtjük. Ekkor a (2.12) formulába a $\| \ \|$ kifejezésben szereplő mennyiségek reprezentálhatók egy $P(x)$ hatványsorral, amelynek a konvergenciasugara nagyobb vagy egyenlő, mint $\min(h_{max}, 2\pi)$. Általánosabb feltételek mellett is közelithető az effektivitási függvény, ha $\|A\| \le a$. A költségfüggvény közelitését is előállitják. A $H(\varepsilon)$ lépéshossz az alábbi egyenlet egyetlen gyöke (lásd (2.5) formula).

$$b(h) \max_{\|A\| \, \|y_o\| \le 1} \| R(hA)(hA)^p (Ay_o) \| = \varepsilon.$$

Sedqwick lemmáját fehasználva kapjuk, hogy

$$\max_{\|A\| \, \|y_o\| \le 1} \| R(hA)(hA)^p (Ay_o) = h^p R^*(h),$$

ahol $R(x) = \sum_{i=0}^{N} r_i x^i$ esetén $R^*(x) = \sum_{i=0}^{N} |r_i| x_i$.

Igy $H(\varepsilon)$ az alábbi egyenlet egyetlen gyöke

$$hP \ b(h) \ R^*(h) = \varepsilon.$$

Végül a költségfüggvényre a (2.7) formulát kapja. A költségfüggvény szigoruan csökkenő.

A 3. ábrán közöljük a vizsgált módszerek költségfüggvényeit, amelyek nem használnak lokális extrapolációt, a 4. ábra pedig a lokális extrapolációt használó módszerek költségfüggvényeinek görbéit ábrázolja.

Néhány megfigyelésüket megemlitjük.

1. Ha a hibakorlát viszonylag erős, a magasabbrendü formulát használó módszer kevésbé költséges, mint az alacsonyabb rendü formulán alapuló. Gyakran előfordul viszont, hogy az alacsonyabbrendü formulák enyhébb hibakorlátok esetén kevésbé költségesek.

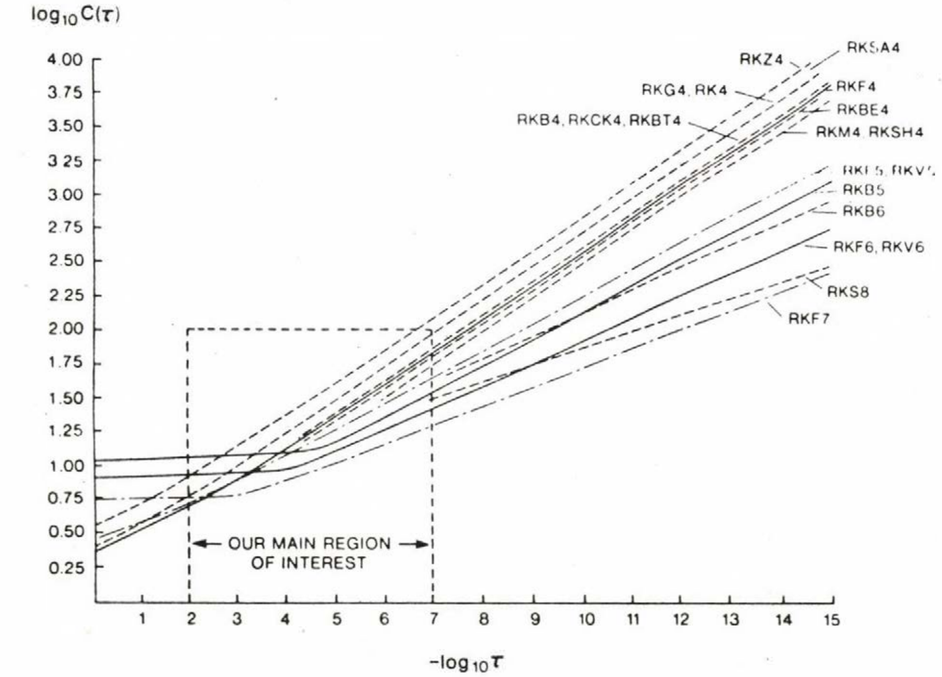
2. A lokális extrapoláció használata lehetőséget nyujt egy módszer viselkedésének javitására. A lokális extrapoláció megnöveli a formula rendjét. Igy elég erős hibakorlátok esetén az extrapoláció bizonyára segit. Ugyanakkor az általuk vizsgált módszerek esetén azok, amelyek lokális extrapolációt használtak, kevésbé költségesek a különböző hibakorlátokra.

3. Általában nem érvényes az a feltételezés, hogy ha az egyik módszer a lokális extrapoláció használata nélkül jobb, mint a másik, akkor abban az esetben is jobb lesz, ha lokális extrapolációt használnak.

4. Általánosan elfogadott nézet, hogy a formula párokon alapuló módszerek jobbak, mint az egy formulából és lépésfelezési hibabecslésből álló módszerek. Az elméleti eredmények ezt igazolni látszanak.

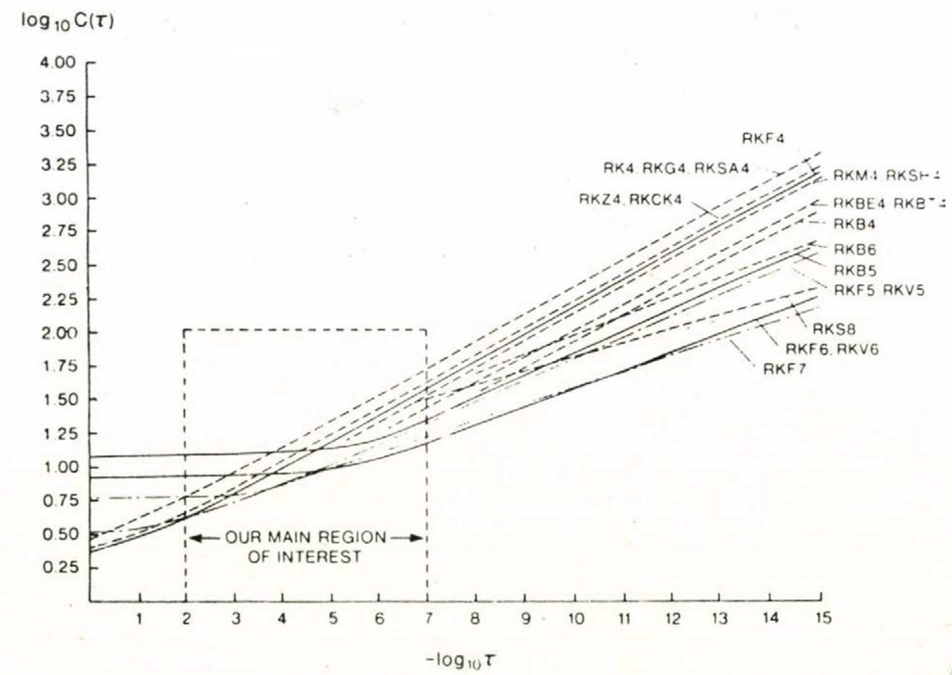
Az adott problémaosztályra a legjobbnak az ötöd- és hatodrendü Verner formulák bizonyultak. (Verner (1978)).

Számitógépes kisérleti eredményeik igazolni látszanak, azt a megállapitást, hogy az általuk konstruált elméleti költségfüggvények az adott problémaosztályra az egyes Runge-Kutta formuláknak és a formulák hibabecsléseinek jó jellemzését adják. Az elméleti módszer ugyanakkor gyorsabb és kevésbé költséges, mint a kisérleti tesztelések.



3. ábra

Költségfüggvénygörbék lokális extrapolációt nem használó módszerekre



4. ábra

Költségfüggvénygörbék lokális extrapolációt használó módszerekre

4. ÖSSZEFOGLALÁS

Az ismertetett elméleti és kisérleti vizsgálatok a módszerek objektiv kiválasztásának lehetséges közelitéseit adják. Érdemük, hogy a numerikus matematikában meghonosodott korábbi gyakorlattal szemben a vizsgálatok feladat és módszer osztályokra vonatkoznak, valamint komplex jellegüek, amennyiben sokféle hatást (hibabecslés, lépéshosszválasztás stb.) vesznek figyelembe. Nyilvánvaló hiányosságuk azonban, hogy a vizsgált feladatosztályok rendkivül speciálisak, és előfordulhat az is, hogy nem azonos rendü módszereket hasonlitanak össze.

Igy általánosnak mondott következtetéseik csak a vizsgált osztályok vonatkozásában fogadhatók el. Eredményeik azonban a korábbi állapothoz képest nagymértékben hozzájárulnak a módszerek megalapozottabb kiválasztásához, noha vizsgálataik következtetései számos esetben ellentmondanak az aszimptotikus vizsgálatok eredményeinek. Ennek oka az, hogy a vizsgálatokat olyan pontossági tartományokra végezték, amelyek az adott konkrét feladattipusok gyakorlati megoldásában előfordulnak.

Megjegyezzük, hogy a fejezetben ismertetett módszerek közül számosnak adaptáltuk a programját az Akadémia CDC 3300-as számitógépére. Többek között ezek a következők:

1. a negyedrendü klasszikus Runge-Kutta módszer
2. Runge-Kutta-Gill módszer
3. Zonneveld módszer
4. Merson módszer
5. Bulirsch-Stoer féle extrapolációs eljárás
6. Fehlberg negyed-ötödrendü módszere
7. Gear módszer stb.

Az algoritmusok és programok rövid leirása a dolgozat 1. táblázatában ill. a CDC 3300-as számitógép programkönyvtári anyagában található (lásd. Numerikus módszerek programgyüjteménye [28]).

III. ÁLTALÁNOSITOTT EGYLÉPÉSES MÓDSZEREK KONVERGENCIÁJA ÉS STABILITÁSA

1. BEVEZETÉS

A fejezetben változó lépéshosszu és változó formulájunak tekinthető általánositott egylépéses módszerekkel foglalkozunk.

Az eredmények egyszerüen adódnak az (1.1), (1.2) közönséges DE-knél szélesebb differenciálegyenletosztályra, funkcionál DE-kre, ezért azokat ilyen általánosabb alakban irjuk le.

Tekintsük az alábbi funkcionál DE-re vonatkozó kezdetiérték problémát. Legyen

$$(3.1)\qquad \begin{aligned} z'(x) &= f(x,z_x), &\quad (0<x<X)\\ z_o(x) &= \varphi(x), &\quad (-\bar{x}\le x\le 0) \end{aligned}$$

$\varphi(x)\in C([-\bar{x},0],R^n)$, ahol Hale (1971) jelöléseit használtuk, azaz $z_x(\theta) = z(x+\theta)(-\bar{x}\le\theta\le 0)$ és $z_x\in C([-\bar{x},0],R^n)$.

A továbbiakban vezessük be a $C_n[x_1,x_2]$ jelölést, amely valós x_1,x_2 értékekre, ha $x_1 \le x_2$, akkor az $[x_1,x_2]$ intervallumon folytonos függvények leképezése R^n-be.

Vizsgáljuk meg most az alábbi kezdeti érték problémát

$$(3.2)\qquad \begin{aligned} y'(x) &= F(x,y), &\quad (a<x\le b)\\ y(x) &= g(x), &\quad (\alpha\le x\le a) \end{aligned}$$

ahol $g\in C_n[\alpha,a]$ kezdő függvény és $F:[a,b]\times C_n[\alpha,b]\to R^n$ olyan, hogy

a/ rögzitett y esetén $F[x,y]$ folytonos minden $x\in[a,b]$

b/ és kielégiti a Lipschitz feltételt y szerint, azaz

$$\| F(x,y_1)-F(x,y_2)\| \le L\| y_1-y_2\|^{[\alpha,x]}$$

minden $y_1,y_2\in C_n[\alpha,b]$ és $x\in[a,b]$.esetén, ahol $\| \;\; \|$ bármelyik természetes R^n norma és

$$\| y_1 - y_2 \|^{[\alpha,x]} = \max_{\alpha \leq s \leq x} \| y_1(s) - y_2(s) \| .$$

A fenti feltételek mellett a (3.2) funkcionálegyenletnek létezik egyértelmü megoldása (Driver (1962)). A későbbiekben mindig feltesszük, hogy ezek a feltételek teljesülnek.

Könnyen belátható, hogy az $x=b-a$, $\bar{x}=b-\alpha$ helyettesitést alkalmazva

$$\varphi(x) = \begin{cases} g(a+x), & (-\alpha \leq x \leq 0) \\ g(a-\alpha), & (-\bar{x} \leq x < -\alpha) \end{cases}$$

$$z(x) = \begin{cases} \varphi(x), & (-\bar{x} \leq x \leq 0) \\ y(x+a), & (0 < x \leq x) \end{cases}$$

és $f(x,z_x) = F(x,y)$, akkor z kielégiti a (3.1) egyenletet és $f:[0,x] \times C_n[-\bar{x},0]$ egy olyan folytonos leképezés, amelyre

$$\| f(x,u) - f(x,w) \| \leq L \| u-w \|$$

minden $x \in [0,X]$ és $u,w \in C_n[-\bar{x},0]$.

Igy beláttuk, hogy a (3.1) és (3.2) egymással ekvivalens feladatok. A továbbiakban a (3.2) jelölési módot fogjuk használni.

Minthogy az $F(x,y)$ funkcionál y szerint kielégiti a Lipschitz feltételt, igy $F(x,y)$ független lesz az $y(s)$ jövőbeli értékeitől, azaz az $s>x$ értékektől. Igy F un. Volterra funkcionál.

A gyakorlatban számos helyen fordulnak elő ilyen funkcionál DE-k. Röviden megemlitjük néhány speciális esetüket. Pl.

a/ a közönséges DE

b/ a retardált DE:

1/ $\left.\begin{array}{ll} z'(x) = f(x,z(x),z(x-\tau(x))), & (0<x<X) \\ z_0 = \varphi\in C_n[-\bar{x},0]. & \end{array}\right\}$

2/ vagy ennek egy rész esete, ha a késleltetés konstans

$\left.\begin{array}{ll} z'(x) = f(x,z(x),\ z(x-\bar{x})), & (0<x<X) \\ z_0 = \varphi\in C_n[-\bar{x},0], & \end{array}\right\}$

ahol $\bar{x}>0$.

3/ egy másik részesete, ha f független $z(x)$-től

$\left.\begin{array}{ll} z'(x) = f(x,z(x-\bar{x})), & (0<x<X) \\ z_0 = \varphi\in C_n[-\bar{x},0], & \end{array}\right\}$ stb.

ahol $\bar{x}>0$.

c/ a Volterra tipusu integro-differenciálegyenletek

$$z'(x) = f(z(x),\ \int_a^x K[y(s),s,x]ds,x]$$

$$z(0) = \varphi(x)\in R^n$$

ahol a K magfüggvény folytonos és egyenletesen kielégiti a Lipschitz feltételt az s és az $y(s)$ szerint.

2. ÁLTALÁNOSITOTT EGYLÉPÉSES MÓDSZEREK

A (3.2) tipusu Volterra funkcionál differenciálegyenletek numerikus megoldási problémáival többen foglalkoztak. Itt most csak az egylépéses módszerekre vonatkozó eddigi eredményeket emlitjük meg.

Az Euler módszer konvergenciáját funkcionál DE-kre Cryer, Tavernini (1972) vizsgálta. Tavernini (1971) általánositott egylépéses módszerekre bizonyitott konvergencia tételt állandó h lépéshossz mellett.

A továbbiakban a fenti eredményt általánositjuk változó lépéshossz esetén.

Definiáljunk a (3.2) egyenletre az [(1.3),k=1] egylépéses módszerek természetes általánositásával kapott közelitő módszert, vezessük be az ilyen módszerek konzisztenciájának, konvergenciájának és keplethibájának fogalmát.

Tekintsük először a legegyszerübb esetet az általánositott Euler módszert. Jelölje $\tilde{y}$ a (3.2) közelitő megoldását. Ekkor

$$\tilde{y}(x_i+rh_i) = \tilde{y}(x_i)+rh_iF(x_i,\tilde{y}),$$
$$(x_i\in\{\Delta_N\}, \quad r\in[0,1], \quad 0\leq i\leq N)$$
$$\tilde{y}(x) = \tilde{g}(x), \qquad (x\in[\alpha,a]).$$

Ezt átirhatjuk a következő ekvivalens alakra

$$\tilde{y}(x_{i+1}) = \tilde{y}(x_i)+h_iF(x_i,\tilde{y}), \quad (x_i\in\{\Delta_N\},r=1),$$
$$\tilde{y}(x_i+rh_i) = r\tilde{y}(x_{i+1})+(1-r)\tilde{y}(x_i), \quad (0<r<1).$$

Igy látható, hogy az $\tilde{y}$ közelitő megoldás értékét az osztópontokban a szokásos Euler módszerrel kapjuk és a többi pontban pedig lineáris interpolációval.

Megjegyezzük, hogy mivel az F értelmezési tartománya $[a,b]\times C_n[\alpha,b]$, igy a közelitő megoldást az $[\alpha,b]$ intervallum minden pontjában ki kell számitanunk, nemcsak az osztópontokban.

Vezessük be most az Euler módszernél általánosabb összefüggést.

3.1 Definició <u>Általánositott egylépéses módszernek</u> nevezzük a következőt

$$\tilde{y}(x_i+rh_i) = \tilde{y}(x_i)+rh_i\emptyset(x_i,\tilde{y},h_i,r),$$
$$(3.3) \qquad (x_i\in\{\Delta_N\}, \quad r\in[0,1], \quad 0<i\leq N)$$
$$\tilde{y}(x) = \tilde{g}(x), \qquad (x\in[\alpha,a]),$$

ahol a $\emptyset$: $S\times C_n[\alpha,b]\times[0,1]\to R^n$ funkcionál a módszer növekményfüggvénye, és $S=\{(x,h_i),\ a\leq x\leq b$ és $0<h_i\leq b-x\leq h^*\}$ a $\tilde{g}$ folytonos függvény pedig legyen g egy közelitése.

Továbbá mindig feltesszük, hogy

a/ $\emptyset(x,y(x),h,r)$ folytonos r szerint ($r\in[0,1]$) minden rögzitett x,y,h esetén

b/ és kielégiti a Lipschitz feltételt, azaz

$$(3.4)\qquad \|\emptyset(x,y_1,h,r)-\emptyset(x,y_2,h,r)\| \leq K\|y_1-y_2\|^{[\alpha,x]}$$

minden $y_1,y_2\in C_n[\alpha,b]$, ($x,h\in S, r\in[0,1]$) esetén, $K>0$ konstans.

A közelitő megoldást itt a korábbiaktól eltérően szükségesnek láttuk megkülönböztetni és az $\tilde{y}$ jelölést bevezetni. Ezzel is hangsulyozni kivánjuk, hogy ez más, függvényjellegü közelités lesz, nemcsak diszkrét pontokban (az osztópontokban) definiálja a (3.2) DE közelitő megoldását, hanem a köztük lévő pontokban is.

A (3.2) funkcionál DE-k (3.3) módszerrel történő megoldásakor definiáljuk az alábbi fogalmakat. Jelölje a (3.2) DE pontos megoldását y.

3.2 Definició A (3.3) általánositott egylépéses módszert <u>konvergensnek</u> nevezzük, ha tetszőleges $\{\Delta_N\}_{N=1}^{\infty}\subset\pi$ felosztássorozat esetén

$$\|y-\tilde{y}\|\to 0, \qquad (\|\Delta_N\|\to 0, \quad \|g-\tilde{g}\|\to 0).$$

3.3 Definició A (3.3) módszer <u>képlethibája</u> az x_i+rh_i helyen

$$\tau(x_i,h_i,r) := y(x_i)+h_i\emptyset(x_i y,h_i,r)-y(x_i+rh_i).$$

Tegyük most fel, hogy a $\emptyset$ olyan, amelyre a képlethiba kielégiti az alábbi feltételt

$$|\tau(x_i,h_i,r)| \leq h_i\varepsilon(x,r,h),$$

ahol ε x,r,h egy megadott hibakorlát függvény, amely függ az x,r és h értékétől.

3.4 Definició A (3.3) módszert <u>konzisztens</u>nek nevezzük, ha

$$\emptyset(x,y,h,1)\to F(x,y) \qquad (\|\Delta_N\| \to 0)$$

egyenletesen x szerint, vagy (ekvivalens)

$$\varepsilon(x,1,h)\to 0 \qquad (\|\Delta_N\| \to 0)$$

minden $x\in[a,b]$ esetén.

A konzisztencia fogalmának talán természetesebb általánositásának tünne, ha azt tennénk fel, hogy

$$\varepsilon(x,r,h)\to 0 \qquad (\|\Delta_N\| \to 0)$$

minden $r\in[0,1]$ értékre; de a (3.3) formulából nyilvánvaló, hogy az $\tilde{y}$ kiszámitásában elkövetett hiba az x_i osztópontoknál egyszerüen összegződik, de az intervallum többi pontjában előbb beszorződik az rh_i tényezővel.

3.5 Definició A (3.3) módszer <u>globális hibájá</u>nak nevezzük az

$$\|y-\tilde{y}\|^{[\alpha,b]}$$

különbséget.

Vezessük be a következő jelölést

$$\nabla(x,\delta) = \begin{cases} \frac{1}{\delta}(y(x+\delta)-y(x)], & \delta>0, \\ y'(x), & \delta=0, \end{cases}$$

minden $(x,\delta)\in S$.

3.6 Definició A (3.3) módszert <u>stabilis</u>nak nevezzük, ha léteznek olyan $c,h^*>0$ konstansok, hogy

$$\|y^*-\tilde{y}\|^{[\alpha,b]} \leq C\ \{g^*-\tilde{g}\| + \sum_{i=1}^{N} \|\delta_i\|\}$$

hacsak $\|\Delta_N\| \leq h^*$, y^* pedig az alábbi perturbált rekurzió megoldása

(3.3') $\quad y^*(x_i+rh_i) = y^*(x_i)+rh_i[\emptyset(x_i,y^*,h_i,r)+\delta_{i+1}]$.

3. AZ ÁLTALÁNOSITOTT EGYLÉPÉSES MÓDSZEREK KONVERGENCIÁJA ÉS STABILITÁSA

Szükségünk lesz a következő tételre.

3.1 Tétel Ha

a/ a $\emptyset$ növekményfüggvény olyan, hogy a képlethibára teljesül

$$\|\emptyset(x,y,h,r)-\nabla(x,r,h)\| \leq \varepsilon(x,h,r) \tag{3.5}$$

minden $x\in\{\Delta_N\}$, $r\in[0,1]$ és $0<h\leq h^*$ esetén,

b/ ahol az ε h-ban monoton függvény és

$$\varepsilon(x,h,r)\leq\bar{\varepsilon}(x,h^*,r)\leq\varepsilon_1 \qquad (\varepsilon_1 \text{ konstans}) \tag{3.6}$$

c/ valamint

$$\bar{\varepsilon}(x,h^*,1)\to 0, \qquad (\|\Delta_N\|\to 0) \tag{3.7}$$

akkor igaz az $[\alpha,b]$ intervallumon a következő becslés:

$$\|y-\tilde{y}\|^{[\alpha,b]}\leq[\|g-\tilde{a}\|^{[\alpha,g]}+h^*\varepsilon_1+\int_a^b\bar{\varepsilon}(x,h^*,1)dx])$$

$$(1+(b-a)Ke^{K(b-a)}) \tag{3.8}$$

<u>Bizonyitás</u> A (3.3) formula szerint $r=1$ esetén

$$\tilde{y}(x_i) = \tilde{g}(a)+\sum_{j=0}^{i-1} h_j\Phi(x_j,\tilde{y},h_j,1)$$

Az $y(x+h) = y(x)+h\nabla(x,h)$ azonosságból

$$y(x_i) = g(a)+\sum_{j=0}^{i=1} h_j\nabla(x_j,h_j).$$

Legyen $\eta = y - \tilde{y}$. A fenti két összefüggést kivonva és a háromszög egyenlőtlenségét alkalmazva kapjuk

$$\| y(x_i) - \tilde{y}(x_i) \| \leq \| g(a) - \tilde{g}(a) \| +$$

$$+ \| \sum_{j=0}^{i=1} h_j [\Phi(x_j, \tilde{y}, h_j, 1) - \nabla(x_j, h_j)] \| \leq$$

$$\leq \| \eta \|^{[\alpha, a]} + \sum_{j=0}^{i-1} h_j \| \Phi(x_j, y, h_j, 1) - \nabla(x_j, h_j) \| +$$

$$+ \sum_{j=0}^{i-1} h_j \| \emptyset\, x_j, \tilde{y}, h_j, 1) - \Phi(x_j, y, h_j, 1) \|$$

A (3.4) Lipschitz feltétel és (3.5) feltétel miatt:

$$\| \eta(x_i) \| \leq \| \eta \|^{[\alpha, a]} + \sum_{j=0}^{i-1} h_j \varepsilon(x_j, h_j, 1) +$$

$$+ k \sum_{j=0}^{i-1} h_j \quad \| \eta \|^{[\alpha, x_j]}$$

A tétel bizonyitásához szükségünk lesz Gronwall-Bellman lemma alábbi diszkrét alakjára (Babuska (1966), Schmidt (1976))

<u>Gronwall-Bellman lemma</u>. Legyenek $x(i)$, $y(i)$, $z(i)$ az $i=0,1,\ldots N$ számokon értelmezett valós függvények és $z(i) \geq 0$ $(i=0,1,\ldots,N)$. Ha az $x(i)$, $y(i)$, $z(i)$ függvények kielégitik az

$$(3.9) \qquad x(i) \leq y(i) + \sum_{j=0}^{i-1} z(j) x(j) \qquad (i=0,1,\ldots,n)$$

egyenlőtlenséget $(n \leq N)$ akkor igaz, hogy

$$(3.10) \qquad x(n) \leq y(n) + \sum_{i=0}^{n-1} z(i) y(i) \prod_{j=i+1}^{n-1} [1 + z(j)].$$

Vonjuk ki most a (3.3) formulát az $y(x+rh) = y(x) + rh\nabla(x, rh)$ azonosságból az $x = x_i$ pontban, hasonlóképpen kapjuk, hogy

$$\| \eta(x_i + rh_i) \| \leq \| \eta(x_i) \| + h_i K \| \eta \|^{[\alpha, x_i]} + h_i \varepsilon(x_i, h_i, r)$$

Irjuk be ide az $\eta(x_i)$ helyébe az előbb kapott korlátot

$$\|\eta\|^{[x_i,x_{i+1}]}\leq\|\eta\|^{[\alpha,a]}+h_i\varepsilon(x_i,h_i,r)+$$

$$+\sum_{j=0}^{i-1}h_j\varepsilon(x_j,h_j,1)+K\sum_{j=0}^{i}h_j\|\eta\|^{[\alpha,x_j]}=P_{i+1}$$

Ugyanakkor

$$\|\eta\|^{[\alpha,x_i]}\leq\max\{\|\eta\|^{[\alpha,a]},P_1,\dots,P_i\}$$ és mivel

$$\|\eta\|^{[\alpha,a]}\leq P_1\leq,\dots,P_i,$$ igy $\|\eta\|^{[\alpha,x_i]}\leq P_i$

Alkalmazzuk a Gronwall-Bellman lemmát a következő szereposztással

$$x(i)=\|\eta\|^{[\alpha,x_i]},\qquad z(i)=Kh_i,$$

$$y(i)=\|\eta\|^{[\alpha,a]}+h_{i-1}\varepsilon(x_{i-1},h_{i-1},r)+\sum_{j=0}^{i-2}h_j\varepsilon(x_j,h_j,1),$$

akkor kapjuk, hogy

$$\|\eta\|^{[\alpha,x_N]}\leq\|\eta\|^{[\alpha,a]}+h_{N-1}\varepsilon(x_{N-1},h_{N-1},r)+$$

$$+\sum_{j=0}^{N-2}h_j\varepsilon(x_j,h_j,1)+\sum_{i=0}^{N-1}Kh_i[\|\eta\|^{[\alpha,a]}+h_{N-1}\varepsilon(x_{N-1},h_{N-1},r)+$$

$$+\sum_{j=0}^{N-2}h_j\varepsilon(x_j,h_j,1)*\prod_{j=i+1}^{N-1}(1+Kh_j)$$

Az $1+x\leq e^x$ egyenlőtlenség miatt

$$\prod_{j=i+1}^{N-1}(1+Kh_j)\leq\prod_{j=i+1}^{N-1}e^{Kh_j}\leq e^{K(b-a)}:$$

Továbbá a (3.6) monotonitási feltételt felhasználva és azt, hogy létezik $h^*=\max_i h_i$

$$\|\eta\|^{[\alpha,b]} \leq \|\eta\|^{[\alpha,a]} + h^{*}\varepsilon_1 + \sum_{j=0}^{N-2} h{*}\bar{\varepsilon}(x,h^{*},1) +$$

$$+ \sum_{i=0}^{N-1} Kh_i e^{K(b-a)} [\|\eta\|^{[\alpha,a]} + h^{*}\varepsilon_1 + \sum_{j=0}^{N-2} h^{*}\bar{\varepsilon}(x,h^{*},1)]$$

Összevonva a megfelelő tagokat és kihasználva, hogy létezik és korlátos az $\int_a^b \bar{\varepsilon}(x,h^{*},1)dx$ Lebesque integrál, kapjuk az alábbi becslést

$$\|\eta\|^{[\alpha,b]} \leq (1+(b-a)Ke^{K(b-a)})[\|\eta\|^{[\alpha,a]} + h^{*}\varepsilon_1 +$$

$$+ \int_a^b \bar{\varepsilon}(x,h^{*},1)dx]$$

Q.E.D.

A fenti becslés segitségével egyszerüen belátható a következő tétel.

3.2 Tétel Ha a (3.3) módszer

a/ konzisztens,

b/ a $\emptyset$ növekmény függvény egyenletesen korlátos x és r szerint, minden $h^{*}\to 0$ esetén azaz $\|\emptyset(x,y,h^{*},r)\| \leq M$, (ahol M konstans)

akkor a (3.3) módszer konvergens minden $\{\Delta_N\}_{N=1}^{\infty}$ felosztássorozatra, ha $\|\Delta_N\| = \max\limits_{0<i\leq N-1} h_i = h^{*}\to 0$ és $\|\tilde{g}-g\| \to 0$.

Bizonyitás

Vizsgáljuk meg, hogy a módszer konzisztenciájából és a $\emptyset$ egyenletes korlátosságából hogyan következnek az előző tétel feltételei. Azaz becsüljük az alábbi mennyiséget

$$\|\emptyset(x,y,h,r)-\nabla(x,rh)\| \leq \|\emptyset(x,y,h,r) -$$

$$- \emptyset(x,y,h,1)\| + \|\emptyset(x,y,h,1)-\nabla(x,rh)\|$$

$$\leq \begin{cases} 2M & r\in[0,1) \\ \bar{\varepsilon}(h^{*})\to 0 & r=1 \end{cases} \quad (h^{*}\to 0)$$

Ezért alkalmazható a (3.8) becslés, azaz

$$\|y-\tilde{y}\|^{[\alpha,b]} \leq [\|g-\tilde{g}\|^{[\alpha,a]}+h^{*}2M+\int_a^b \bar{\varepsilon}(x,h^{*},1)dx] *$$

$$*(1+(b-a)Ke^{K(b-a)})$$

Igy

$$\|y-\tilde{y}\|^{[\alpha,b]}\to 0 \text{ , amikor } h^{*}\to 0 \text{ és } g\to\tilde{g} \quad \text{Q.E.D.}$$

3.3 Tétel A (3.3) alaku egylépéses módszerek stabilisek.

Bizonyitás

A (3.3) és (3.3') formula szerint $r=1$ esetén kapjuk, hogy

$$y^{*}(x_i)=g^{*}(a)+\sum_{j=0}^{i-1} h_j[\Phi(x_j,y^{*},h_j,1)+\delta_{j+1}]$$

$$\tilde{y}(x_i)=\tilde{g}(a)+\sum_{j=0}^{i-1} h_j\Phi(x_j,\tilde{y},h_j,1).$$

Legyen $\eta=\tilde{y}-y^{*}$. Vonjuk ki egymásból a fenti két egyenlőséget és használjuk fel a háromszög egyenlőtlenséget és azt, hogy a $\emptyset$ kielégiti a Lipschitz feltételt y szerint. Akkor

$$\|\eta(x_i)\| = \|y^{*}(x_i)-\tilde{y}(x_i)\| \leq \|g^{*}(a)-\tilde{g}(a)\| + \sum_{j=0}^{i-1} h_j\|\delta_{j+1}\| +$$

$$+\sum_{j=0}^{i-1} h_j\|\Phi(x_j,y^{*},h_j,1)-\Phi(x_j,\tilde{y},h_j,1)\| \leq$$

$$\leq \|\eta\|^{[\alpha,a]}+\sum_{j=0}^{i-1} h_j\|\delta_{j+1}\| + \sum_{j=0}^{i-1} h_j K\|y^{*}-\tilde{y}\|^{[\alpha,x_j]}$$

Vonjuk ki most a (3.3) és (3.3') formulát egymásból

$$\|\eta\|^{[x_i,x_i+1]} = \|y^{*}(x_i+rh_i)-\tilde{y}(x_i+rh_i)\| \leq \|y^{*}(x_i)-\tilde{y}(x_i)\| +$$

$$+h_i\|\Phi(x_i,\tilde{y},h_i,r)-\Phi(x_i,y^{*},h_i,r)\| + h_i\|\delta_{i+1}\| \leq$$

$$\leq \|\eta\|^{[\alpha,a]}+K\sum_{j=0}^{i} h_i\|\tilde{y}-y^{*}\|^{[\alpha,x_j]}+\sum_{j=0}^{i} h_j\|\delta_{j+1}\| = P_{i+1}$$

$$\|\eta\|^{[\alpha,x_i]} \le \max\{\|\eta\|^{[\alpha,a]}, P_i\}$$ és mivel

$$\|\eta\|^{[\alpha,a]} \le P_1 \le \dots \le P_i,$$ igy $\|\eta\|^{[\alpha,x_i]} \le P_i$.

A Gronwall-Bellmann lemmát alkalmazva az alábbi szereposztásban

$$x(i) = \|\eta\|^{[\alpha,x_i]}, \qquad z(i) = Kh_i,$$

$$y(i) = \|\eta\|^{[\alpha,a]} + \sum_{j=0}^{i} h_j \|\delta_{j+1}\|$$

kapjuk, hogy

$$\|\eta\|^{[\alpha,b]} \le \|\eta\|^{[\alpha,a]} + \sum_{j=0}^{N} h_j \|\delta_{j+1}\| + \sum_{i=0}^{N-1} Kh_i [\|\eta\|^{[\alpha,a]} +$$

$$+ \sum_{j=0}^{i} h_j \| \quad_{j+1}\|] \cdot \prod_{j=i+1}^{N-1} (1+Khj)$$

Az $1+x \le e^x$ összefüggés miatt:

$$\prod_{j=i+1}^{N-1} (1+Kh_j) \le \prod_{j=i+1}^{N-1} e^{Kh_j} \le e^{K(b-a)}.$$

Igy

$$\|\eta\|^{[\alpha,b]} \le \|\eta\|^{[\alpha,a]} + h^* \sum_{j=0}^{N} \|\delta_{j+1}\| + Ke^{K(b-a)} \sum_{i=0}^{N-1}$$

$$[h_i \|\eta\|^{[\alpha,a]} + \sum_{j=0}^{i} h_j \|\delta_{j+1}\|]$$

$$\le \|\eta\|^{[\alpha,a]} (1+K(b-a)e^{K(b-a)}) + \sum_{j=0}^{N} \|\delta_{j+1}\| (h^* + Kh^* e^{K(b-a)})$$

$$\le (C_1 + K(b-a)e^{K(b-a)}) (\|\eta\|^{[\alpha,a]} + \sum_{j=0}^{N} \|\delta_{j+1}\|).$$

ahol $C_1 = \max(1, h^*)$. Q.E.D.

Módszer (program neve)	FORMULA	HIBABECSLÉS	Döntési kritérium (lépésfogadásának feltétele) lépéshossz (rend) változtatási mechanizmus
Zonneveld (1964)	ötödrendü RK: $y_{n+1}=y_n+(35k_o+162k_2+125k_4+14k_5)/336$; $k_o=hf(y_n), k_1=hf(y_n+2/9k_o), k_3=hf(y_n+(k_o+3k_2)/8)$, $k_4=hf(y_n+(53k_o-135k_1+126k_2+56k_3)/125$, $k_5=hf(y_n+ -63k_o+189k_1-36k_2-112k_3+50k_4)/28$.	az y_n h szerinti Taylor sorfejtésének 5. tagja $th^5dy=(21k_o-162k_2+224k_3-125k_4+42k_6)/14$; $k_6=hf(y_n+(133k_o-378k_1+276k_2+112k_3+25k_4)/168$	th^5dy (abs(f*relativ hibakorlát+absz.h.k.)*abs(h)/(b-a) 1. számolunk h lépéshosszal; 2. kiszámitjuk a $h_{uj}=0.95h\sqrt[5]{\varepsilon/abs(th^5dy)}$ 3. kétlépéses elfogadásáig 2./ alapján számolunk; 4. kezdetben vagy elutasitás után $h_{uj}=mu_1*h$, $mu_i=től(x_i)/(től(x_i)+abs(th^5dy(x_i))=0.45$; 5. különben $h_{uj}=h((h/h_{előző}+1)mu_1-mu_o$
Gear (1971) DIFSUB (DIFSU,DIFSU2 a CDC-n)	nem stiff DE-re: Adams-Basforth-Moulton (1.10) azonos rendü prediktor-korrektor pár (1-13-ad rendig). stiff DE-re: (BDF) retrográd differenciálás formula (1-6-ad rendig)	főhibatag sulyozott becslése osztott differenciákkal (p-1)-ed rendre: $c_p p!\lVert a_p/\omega\rVert_2=\alpha_{p-1}$, mivel $a_p\approx\frac{h^P}{P!}y^P$ p-ed rendre: $C_{p+1}!\lVert\nabla a_p\omega\rVert_2=\alpha p$, (p+1)-ed rendre: $C_{p+2}P!\lVert\nabla^2_{a_p}p\omega\rVert_2=_{p+1}$, mivel $\nabla^2_{a_p}=\frac{h^{P+2}}{P!}y^{p+2}$, $a=[y,hy',\dots,h^Py^{(P)}/p!]^T$, ∇a_p az utolsó komponens retrográd differenciája; ω sulykomponens, $\lVert\ \rVert_2$ L_2 norma	$\alpha_p\le\varepsilon$ (megadott hibakorlát). lok.opt. lépéshossz: $c_{p+1}p!\alpha^{p+1}\lVert\nabla a_p/\omega\rVert_2=\varepsilon$; $h_{uj}=0.99\ \alpha h$, ahol $\alpha=\frac{1}{1.2}[\frac{\varepsilon}{\alpha}]^{1/p+1}$ p-ed rendre; $\alpha=\frac{1}{1.4}[\frac{\varepsilon}{\alpha_{p-1}}]^{1/p}$ (p-1)-ed rendre, $\alpha=\frac{1}{1.3}[\frac{\varepsilon}{\alpha_{p+1}}]^{1/p+2}$ (p+1)-ed rendre; α-t számoljuk, ha: 1.a lépés sikertelen volt, kivéve, ha éppen a rendet akartuk növelni; 2. (p+1) lépéssel az utolsó lépéshossz, ill. rendváltoztatás után; 3. tiz lépéssel az utolsó α becslés után, hacsak közben nem növeltük a lépéshosszt.a legnagyobb α-nak megfelelő rendü módszert választjuk, αh is változik.
Lindberg (1972) IMPEX (STIFF-CDC)	Implicit érintő formula: $y_{n+1}=y_n+hf((y_n+y_{n+1})/2)$, és passziv simitás az $x_i=x_o+ih_o$ (i=1,2,...) pontokban: $(h=h_o,h_o/2)$ $\hat{y}_n(h)=[y_{n-1}(h)+2y_n(h)+y_{n+1}(h)]/4$; passziv extrapoláció $\bar{y}_n(h)=\hat{y}_n(h_o/2)+[\hat{y}_n(h_o/2)-\hat{y}_n(h)]/3$; (lépéshosszváltásnál igy indul)	globális hiba becslése; a lokális hiba becslése lépésfelezéssel és extrapolációval $est=c_n-h^5\simeq 1/36\lVert\Delta^3_n[\{y_n(h/2)-y_n(h)\}3]\rVert$.	$est\le\varepsilon_{Lok}$, $\varepsilon=2\frac{\sqrt{12\varepsilon}}{\lVert x\rVert}e^{\frac{-2\sqrt{12\varepsilon}}{\lVert x\rVert}}$ est/ε_{Lok} — h_{uj} $<1/5110$ — $(\sqrt[5]{\varepsilon_{Lok}/5.est})h$ 1/5110, 1/80 — 2h 1/20, 1/2 — változatlan >1 — az utolsó lépés ujra
Shampine, Watts (1976) RKF45	4-ed és 5-öd rendü Fehlberg form.(1969) $y_{n+1}=y_n+h(16/135k_1+6656/12825k_3+28561/56430k_4-9/50\ k_5+2/55k_6)$ (5-öd); $y_{n+1}=y_n+h(25/216k_1+1408/2565k_3+2197/4104k_4-1/5\ k_5)$; $k_1=f(x_n,y_n)$; $k_2=f(x_n+1/4h, y_n+1/4k_1)$; $k_3=f(x_n+3/8h, y_n+3/32k_1+9/32k_2)$; $k_4=f(x_n+12/13h, y_n+1932/2197*k_1-7200/2197k_2+7296/2197\ k_3)$; $k_5=f(x_n+h, y_n+439/216*k_1-8k_2+3680/513k_3-845/4104k_4)$; $k_6=f(x_n+1/2h; y_n-8/27k_1+2k_2-3544/2565k_3+1859/4104k_4-11/40k_5)$	lokális hiba becslése (beágyazási tipusu módszer) $est=h(-1/360k_1+128/4275k_3+2197/75240k_4-1/50k_5-2/55k_6)$.	$est\le\lvert y\rvert*$ rel.hibakorlát + absz.h.k.$=\varepsilon_{összes}$ (relativ h.k.a lépés kező és végpontjában). $h_{uj}=0.9\sqrt[5]{\frac{\varepsilon}{est}}h$; $h_{min}=0.26*u$ (ahol u az adott számitógépen legkisebb pozitiv szám), sikertelen lépés után: a lépéshossz nem növelhető; az 1/10-nél nem jobban lecsökkentett lépéshosszal megismétli a lépést; teljesülnie kell a $h_{uj}\ge h_{min}$ feltételnek
Wanner (1976) STIFFI	negyedrendü Hermite-Obreschkoff formula $y_{n+1}-h/2\ y'_{n+1}+h^2/6\ y''_{n+1}/2! = y_n+h/2\ y'_n+h^2/6\ y''_n/2!$	főhibatag becslése $est=h\ 5/6y^5/5!$ lépésfelezéssel osztott differenciák segitségével.	$est=\varepsilon$ (adott hibakorlát); $h_{uj}=\sqrt[5]{6\varepsilon/\max_{i,n}\lvert y_i[x_{n-2},x_{n-2},x_{n-1},x_{n-1},x_n,x_n]\rvert}$, ha $est\ge 20\varepsilon$, akkor a h_{uj} lépéshosszal megismétli a lépést; különben fenti formula szerint változtatja h-t, de nem engedi meg, hogy ötszörösénél nagyobbra nőjjön.
Alexander (1977) DIRK	diagonálisan implicit RK-módszer (együtthatómátrix alsó háromszög tipusu azonos diagonális elemekkel) pl. 2-pontos 2-od rendü: $y_{n+1}=y_n+(1-\alpha)k_1+\alpha k_2$; ahol $k_1=hf(x_n+\alpha h, y_n+\alpha k_1)$, $k_2=hf(x_n+h, y_n+(1-\alpha)k_1+\alpha k_2)$; $(\alpha=1\pm\frac{1}{2}\sqrt{2})$.	főhibatag sulyozott becslése az RMS normában $(\lVert y\rVert = 1/m\sum_{i=1}^{m}(y^{(i)}/y^{(i)}_{max})^2)^{1/2}$ m-pontos formulára). $\lVert y_{n+1}-y_{n+2/2}\rVert/(2^P-1)$	ha est>ε a lépést elutasitjuk és h-t csökkentjük ugy, hogy a várt hiba ~ε/5 legyen $3\varepsilon/4<est\le\varepsilon$, a lépést elfogadjuk, de h-t fenti módon lecsökkentjük; $\varepsilon/10<est\le 3\varepsilon/4$, a lépést elfogadja és ugyanazzal számol tovább; ha est ε/10, a lépést elfogadja és h-t megnöveli ugy, hogy a várt hiba ε/2 legyen; feltéve, hogy a/ legalább p+1 sikeres lépést tettünk a h utolsó csökkentése óta; b/ csökkentés után ha legfeljebb kétszeresére nőtt, mialatt a legnagyobb megengedett növekedés 10-szeres; c/ a növekedés legalább 1.3-szoros.
Cash, Liem (1978)	m-pontos semi implicit Rosenbrock formula (1963): $y_{n+1}=y_n+h\sum_{r=1}^{m}c_rk_r$; $k_1=f(y_n)+\alpha_1h\frac{\partial f}{\partial y}(y_n)k_1$; $k_r=f(y_n+h\sum_{s=1}^{r-1}b_{rs}k_s+\alpha_rh\frac{\partial f}{\partial y}(y_n+h\sum_{s=1}^{r-1}\beta_{rs}k_s)k_r$ $(r=2,3,\dots,m)$ $c_i,\alpha_i,\beta_{ij},b_{ij}$ konstans; 3-ad rendü 3-pontos eset.	a lokális hiba főtagjának becslése lépésfelezéssel (est)	a/ $\max\lvert est_{n+2}\rvert_j>\varepsilon\to h_{uj}=h/2$ és megismétli a lépést (x_n,y_n)-től; (j=1,2,...,m egyenletek száma); b/ $\max\lvert est_{n+2}\rvert_j<\varepsilon/25$ $h_{uj}=2h$ lépéshosszal folytatja az (x_{n+2},y_{n+2}) ponttól; c/ különben h rögzitett ßs folytatjuk tovább az (x_{n+2},y_{n+2}) ponttól
Zlatev, Thomsen (1979)	(BDF) retrográd differenciálási formula (harmadrendü)	lokális hiba főtagjának becslése (est)	$125\ est<\varepsilon\to h_{uj}=4h$; $16\ est<\varepsilon\le 125\ est\to h_{uj}=2h$; $1/16\ est<\varepsilon\le 16\ est\to h_{uj}=h$; $\varepsilon\le 1/16\to est\ h_{uj}=h/2$;

1. táblázat

Néhány ismert módszer leirása

IRODALOMJEGYZÉK

1. Alexander R. (1977) Diagonally implicit Runge-Kutta methods for stiff O.D.E.S
SIAM J. Numer.Anal. V.14. N.6. Dec. 1007-1027

2. Babuska I., Prãger M., Vitasek E. (1966) Numerical processes in DE's
Wiley (Interscience) London

3. Bulirsch R. and Stoer J. (1966) Numerical treatment of O.D.E.'s by extrapolation methods
Num. Math. 8. 1-13.

4. Butcher J.C. (1966) On the convergence of numerical solutions to O.D.E.'s
Math. of Comput. 20. 1-10.

5. Cash J.R., Liem C.B. (197) On the computational aspects of semi-implicit Runge-Kutta methods
The Computer Journal V.24. N.4. 363-365

6. Ceschino F. and Kuntzman J. (1966) Numerical solution of Initial Value Problems
Prentice Hall, Englewood Cliffs

7. Coddington, Levinson (1955) Theory of O.D.E.'s
Mc Graw-Hill Book Comp. Inc. New York

8. Dahlquist G. (1956) Convergence and stability in the numerical integration of O.D.E.'s
Math. Scandinavica 4. 33-53

9. Driver R.D. (1962) Existence and stability of solutions of a delay-differential system
Arch. Rational Mech. Anal. 10. 401-426.

10. England,R. (1969) Error estimates for Runge-Kutta type solutions to systems of O.D.E.'s
Computer Journal 12. 166-169.

11. Fehlberg,E. (1969) Klassische Runge-Kutta Formeln fünfter und siebenter Ordnung mit Schrittweiten Kontrolle Computing, 4, 93-106.

12. Galántai A. (1978) Vizsgálatok a közönséges differenciálegyenletek közelitő módszereinek konvergencia és hibaanalizisének körében
Kandidátusi disszertáció

13. Galántai A. - Strehó M. (1977) Stiff problems I. ELTE Numerikus és Gépi Matematikai Tanszék kiadványa
Budapest

14. Gear, C.W. (1971) Numerical Initial Value Problems in O.D.E.'s
Prentice Hall, Englewood Cliffs, New Jersey

15. Gear C.W. Tu K.W. (1974) The effect of variable mesh size on the stability of multistep methods
SIAM J. Numer.Anal. V. 11.N.5. October 1025-1043.

16. Gear C.W., Watanabe D.S. (1974) Stability and convergence of variable order multistep methods
SIAM J. Numer.Anal. V.11. N.5. October 1044-1058.

17. Gill, S. (1951) A process for the step-by-step integration of DE's in an automatic computing machine
Proc. Cambridge Philos. Soc. 47. 96-108.

18. Hahn, W. (1967) Stability of Motion
Springer Verlag

19. Hale J. (1971) Functional Differential Equations
New York, Springer

20. Hall, G., Watt, J.M. (1976) Modern numerical methods for O.D.E.'s
Clarendon Press, Oxford 1976.

21. Henrici, P. (1962) Discrete Variable Methods in Ordinary Differential Equations Wiley New York

22. Hull, T.E. (1969) The numerical integration of O.D.E.'s Information Processing 68 - North-Holland Publ. Comp. Amsterdam 40-53.

23. Hull, T.E., Enright, W.H., Fellen, B.M. and Sedgwick A.E. (1972) Comparing numerical methods for O.D.E.'s SIAM J. Numer. Anal. 11. 965-978.

24. Jackson, K.R., Enright W.H., Hull T.E. (1978) A theoretical criterion for comparing Runge-Kutta formulas
SIAM J. Numer. Anal. V.15. N.3. June 618-641.

25. Krogh, F.T. (1973) Algorithms for changing the step size
SIAM J.Numer.Anal. 10. 949-965.

26. Lindberg,B. (1972) IMPEX-A program package for solution of systems of stiff differential equations
Report NA 72.50, The Poyal Institute of Technology Stockholm Sweden

27. Nordsieck A. (1962) On the numerical Integration of O.D.E.'s
Math. Comp. 16. 22-49.

28. Numerikus módszerek programgyüjteménye I. II. III.
MTA SzTAKI 1971. 1973. 1976.

29. Pontrjagin, Boltyanszkij (1968) Optimális folyamatok elmélete
Közgazdasági és Jogi Könyvkiadó, Budapest

30. Rosenbrock H.H. (1963) Some General Implicit Processes for the Numerical Solution of D.E.'s
Comp. J. 5. 329-330.

31. Sacks-Davis R. (1977) Error Estimates for a stiff Differential Equation Procedure
Mathematics of Computation V.31. N.140. Oct. 939-953.

32. Shintani,H. (1966) Two step processes by one-step methods of order 3 and of order 4.
Ibid 183-195.

33. Sarafyan D. (1968) Estimaticn of errors for the approximate solution of differential equations and their systems,
Riv.Mat.Univ. Parma 9. 109-127.

34. Schmidt, J.W. (1976) Über lineare Ungleichungen vom Gronwallschen Typ
Beiträge zur Num. Math. 5. 171-189.

35. Shampine, L.F. (1973) Local extrapolation in the solution of O.D.E.'s
Mathematics of Computation 27. 91-97.

36. Shampine L.F. and Watts H.A. (1976) Practical Solution of O.D.E.'s by Runge-Kutta methods
SAND 76-0585 U.S.A.

37. Spijker, M.N. (1966) Convergence and stability of step-by-step methods for the numerical colution of initial values problems
Numerische Mathematics 8. 161-177.

38. Stetter, H.J. (1971) Stability of discretizations on infinite intervals
Lecture Notes in Mathematics 228. Springer Verlag 207-222.

39. Stetter, H.J. (1973) Analysis of Discretization Methods for O.D.E's
Springer Verlag

40. Sterhó M. (1976) Stiff tipusu közönséges differenciálegyenletek megoldásáról
SzTAKI Tanulmányok 48.

41. Tavernini L. (1971) One step methods for the numerical solution of Volterra funct. diff. equations
SIAM J. Numer. Anal. 8. (1971) 766-795.

42. Тиман А.Ф. (1960) Теория приближения функции действительного переменного
Гос. издат. физ. мат. лит. Москва

43. Varga R.S. (1961) On higher order stable implicit methods for solving parabolic partial DE's
Journ. Math. and Phys. 40, 220-231.

44. Varga R.S. (1971) Functional analysis and approximation theory in numerical analysis
Regional conference Series in Applied Mathematics, SIAM

45. Wanner G. (1976) STIFFI, A program for O.D.E's
Report, Universite de Geneve Oct.

46. Verner J.H. (1978) Explicit Runge-Kutta methods with estimates of the local truncation error
SIAM J. Numer. Anal. V.15. N.4. August 772-790.

47. Zlatev Z. (1979) Application of backward differentation methods to the finite element solution of time-dependent problems
Intern.J. for Num. Meth. in Engin. V.14. 1051-1061.

48. Zonneveld J.A. (1964) Automatic Numerical Integration
Mathematisch Centrum, Amsterdam.

49. Cryer, C.W., L. Tavernini (1972) The numerical Solution of Volterra functional differential equations by Euler's method
SIAM J. Numer.Anal. 9. 105.-129.

50. Sedgwick A.E. (1973) An effective variable order variable step Adams method
Ph.D. thesis, Dept. of Computer Science Tech. Rep.N.53, University of Toronto, Toronto.

SUMMARY

The first part of the paper (Chapters I. II.) contains the exact description of the problems arising in the field of the numerical solution of initial value problems for ordinary differential equations. The results are based on the critical selection and interpretation of the literary ones. The comparative study enables us to find the appropriate methods for the solution of concrete problems. The computer experiences accumulated in the course of the practical solution of such tasks are described.

The second part (Chapter III) contains the definition of initial value problems of functional differential equations (in which derivatives are dependent on the earlier values of the solution, too) and 2 theorems with proofs established by the author. They are of basic importance in the theoretical study of general one-step methods for the funcitonal differential equations concerning their convergence and stability in the case of variable stepsize.

From the literature 50 references are selected.

The study facilitates the adaptation of the numerical method to tasks of application on the one hand, and its comprehensive character suggests further investigation on the other.

Р Е З Ю М Е

Первая часть этой работы (главы I, II) содержит в себе математически точное описание проблем возникающихся при численном решении задачи Коши для обыкновенных дифференциальных уравнений. Результаты основаны на критическом выборе и интерпретации литературных данных. Сравнительные испытания помогают найти подходящий метод для решения конкретных задач. Опыты собранные на ЭВМ также описаны.

Во второй части (глава III.) дается определение задачи Коши для функциональных дифференциальных уравнений (в которых производные зависят от бывших значений решения также) и высказаны и доказаны 2 теоремы автора. Эти являются очень важными в теоретических исследованиях обобщенных одношаговых методов для решения функциональных дифференциальных уравнений. Доказывается сходимость и устойчивость этих методов в случае, когда шаг изменяется.

Задается из литературы 50 работ.

A TANULMÁNYSOROZATBAN 1980-BAN JELENTEK MEG:

101/1980 Gerencsér László - Hangos Katalin:
Diszkrét lineáris sztochasztikus rendszerek
önhangoló szabályozása

102/1980 Pásztorné Varga Katalin: Rekurziv eljárás

103/1980 Gerencsér Piroska - Szép Endre - Zilahy Ferenc
Marton Zsolt: Robotmegfogók adaptivitása I.

104/1980 Knuth Előd - Radó Péter - Tóth Árpád:
A SDLA előzetes ismertetése

105/1980 E. Knuth - P. Radó - A. Toth:
Preliminary description of SDLA

106/1980 Prékopa András: Sztochasztikus programozási
modellek és alkalmazásuk

107/1980 Kelle Péter: Megbizhatósági készletmodellek és
alkalmazásuk

108/1980 Almásy Gedeon: Mérlegegyenletek és mérési hibák

109/1980 Békéssy A. - Demetrovics J. - Gyepesi Gy.:
Relácios adatbázis logikai szintü vizsgálata
funkcionális függőségek szempontjából

110/1980 Gaál A. - Soltész J. - Ruda M. - Ratkó I.:
Tanulmányok a statisztikai adatfeldolgozásról

111/1980 Benedikt Szvetlána: Nem ismételhető döntéshozatal
analizise kockázattal járó esetekben

112/1980 Verebély Pál: Többprocesszoros, osztott intelligen-
ciáju grafikus rendszerek tervezési és megvalósitási
kérdései

113/1980 V. Visegrádi Téli Iskola

114/1980 Demetrovics János: Relációs adatmodell logikai és strukturális vizsgálata

115/1980 Gergely József: Program package for sparse matrices

1981-BEN JELENTEK MEG:

116/1981 Siegler András: Egy 6 szabadságfoku antropomorf manipulátor kinematikája és számitógépes vezérlése

117/1981 Knuth Előd - Radó Péter: Principles of Computer Aided System Description

118/1981 Demetrovics János - Gyepesi György: Általános függések

119/1981 Sztanó Tamás: REAL-TIME programrendszerek eseményvezérelt szervezése

120/1981 Szentgyörgyi Zsuzsa: A számitástechnika műszaki fejlődése és társadalmi hatásai